ADVANCING THE SAFETY OF LENTIVIRAL VECTOR

MEDIATED GENE THERAPY

Aaron Marcus Shaw

Submitted to the faculty of the University Graduate School

in partial fulfillment of the requirements

for the degree

Doctor of Philosophy

in the Department of Medical and Molecular Genetics

Indiana University

April 2015

Accepted by the Graduate Faculty, of Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Kenneth Cornetta, MD, Chair

_____

Nuria Morral, PhD

Doctoral Committee

_____

Scott R. Witting, PhD

January 7, 2015

_____

Mervin C. Yoder, MD

ACKNOWLEDGEMENTS

I would like to thank my mentor Dr. Ken Cornetta for affording me the opportunity to pursue my interest in the field of gene therapy as a graduate student in his laboratory. Dr. Cornetta allowed me to work independently while providing invaluable guidance to keep me on track with my projects. He placed an emphasis on getting the most out of my time in the classroom as well as the laboratory.

I would also like to thank the other members of my graduate research committee, including Dr. Merv Yoder who provided innovative ideas to help keep my project novel and in line with current research topics. I would like thank Dr. Nuria Morral for her expertise in gene therapy proving invaluable to my research. Lastly, I would like to give a large thank you to Dr. Scott Witting who tolerated my almost daily requests for help and guidance over the years and whose gene therapy knowledge was instrumental in my success.

I would like to acknowledge my fellow lab members past and present for their technical assistance and guidance without which my project never would have gotten off the ground. I would like to thank the Medical and Molecular Genetics department for providing valuable assistance and leadership. I am grateful for the funding that made my work possible, provided in part by the Joe and Shirley Christian Graduate Student Education Fund. Finally, I would like to thank my fellow IBMG students for their support and friendship, my mentors from the U.S. Navy who instilled my sense of leadership and responsibility, and my family for supporting me in all my endeavors.

Aaron Marcus Shaw

ADVANCING THE SAFETY OF LENTIVIRAL VECTOR MEDIATED GENE THERAPY

Lentiviral vector mediated gene therapy has made great strides in recent years with several successful clinical trials. However, adverse events encountered with some early trials have highlighted the necessity to improve upon its safety. Improvements can range from early steps in vector production to evaluation of insertion sites post-transduction. We have evaluated an FDA approved DNase for removal of residual plasmid DNA during vector production, developed novel non-integrating lentiviral vectors and employed modified insertion site analysis post-transduction to improve the safety of lentiviral vector mediated gene therapy.

To prevent the exposure of gene therapy patients to HIV-1 DNA it is essential to remove residual plasmid DNA during vector production. We evaluated a recombinant human DNase which has been FDA approved for use in patients as an alternative to a bacterially derived DNase. Our results indicate this DNase is an effective alternative with a potentially safer profile for use in patients.

The ability of lentiviral vectors to stably integrate their genome into a host cell's DNA can have negative side-effects due to the risk of insertional mutagenesis. Non-integrating lentiviral vectors have been developed to alleviate this risk in applications where integration is not necessary. However, a low frequency of illegitimate integration persists when using these vectors. We have developed a novel non-integrating vector

mutation and evaluated the efficacy of combining it with other mutations for reducing the frequency of illegitimate integration. We demonstrate that combining mutations that inhibit integration can further reduce the frequency of illegitimate integration.

Several methodologies have been developed for evaluating the insertion sites of normal integrating lentiviral vectors. Illegitimate integration by non-integrating vectors demonstrates mechanisms which result in insertions and/or deletions at the vector-genome junction. Current methods lack the sensitivity to account for these variables in a high-throughput manner. We have adapted modifications to current methods to improve the capture of these variable insertion sites for analysis.

The results of these studies improve the safety of lentiviral vector mediated gene therapy by improving the purity of the vector product, providing a safer vector for non-integrase mediated applications, and allowing more sensitive analysis of insertion sites post-transduction.

Kenneth Cornetta, MD, Chair

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

cDNA	Complementary deoxyribonucleic acid
CMV	Cytomegalovirus early promoter
cPPT	central polypurine tract
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DSB	Double strand break
GFP	Green fluorescence protein
HEK	Human embryonic kidney
HIV-1	Human immunodeficiency virus type 1
HR	Homologous recombination
LAM-PCR	Linear amplification mediated polymerase chain reaction
LM-PCR	Ligation mediated polymerase chain reaction
LTR	Long-terminal repeat
LV	Lentiviral vector
MFI	Mean fluorescence intensity
NHEJ	Non-homologous end-joining
NILV	Non-integrating lentiviral vector
nrLAM-PCR	Nonrestrictive linear amplification mediated polymerase chain reaction
PIC	Pre-integration complex
PCR	Polymerase chain reaction

PPT	polypurine tract
Q-PCR	Quantitative-polymerase chain reaction
RNA	Ribonucleic acid
RRE	Rev-response element
RSV	Rous sarcoma virus promoter
SIN-LTR	Self-inactivating LTR
U5/U3	Untranslated 5’ and 3’ regions
VSV-G	Vesicular Stomatitis Virus G glycoprotein
WPRE	Woodchuck Hepatitis Virus Post-Transcriptional Response Element

I. Introduction

A. Design and production of lentiviral vectors

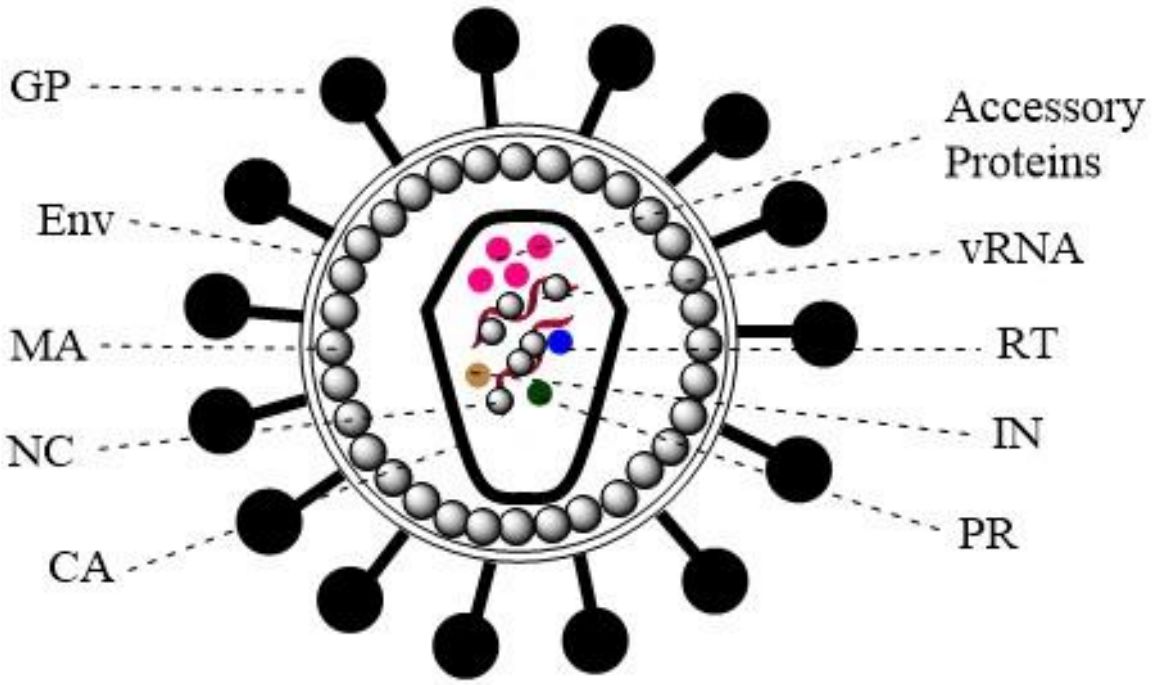
"Adapted from Shaw A, Cornetta K. Design and Potential of Non-Integrating Lentiviral Vectors. Biomedicines. 2014; 2(1):14-35"

Gene therapy using the current third generation of lentiviral vectors (LV) holds great promise for the treatment of a wide variety of disorders. To date, the predominant use of these vectors has been ex vivo modification of hematopoietic stem cells, T cells or other targets where the transduced cell is expected to expand in vivo. For example, early clinical trials have shown evidence of disease correction or delay of onset in the treatment of thalassemia, adrenoleukodystrophy, chronic lymphocytic leukemia, metachromatic leukodystrophy, and Wiskott Aldrich syndrome (Aiuti et al., 2013; Biffi et al., 2013; Cartier et al., 2009; Cavazzana-Calvo et al., 2010; Porter, Levine, Kalos, Bagg, & June, 2011). These accomplishments would not have been possible without the continuous improvements implemented in LV design which have accrued incrementally through several generations to help increase the safety and efficacy for their use in clinical applications.

LV are a subclass of retroviral vectors which are derived from the human immunodeficiency virus type 1 (HIV-1) lentivirus and have the ability to integrate efficiently into quiescent or non-dividing cells (Naldini et al., 1996). As illustrated in

Figure 1, the wild-type lentivirus has an outer membrane or envelope which consists of a lipid bilayer with surface glycoproteins for engaging receptors on its target cells. The inside of the virus consists of an inner matrix and a capsid core. Inside of the capsid are two copies of a single-stranded viral RNA (vRNA) genome complexed with a nucleocapsid for protection from degradation along with the viral proteins necessary for maturation, infection and replication.

Figure 1. Outline of a wild-type lentivirus. Schematic of a wild-type lentivirus as described in the text. GP – surface glycoproteins, Env – envelope or viral membrane, MA – inner matrix, NC – nucleocapsid, CA – capsid, PR – protease, IN – integrase, RT – reverse transcriptase, vRNA – viral RNA genome.



The life-cycle of a lentivirus requires several key steps to maintain an active infection (Figure 2). Briefly, (1) the virus first engages receptors on the target cell with its surface glycoproteins. This leads to a conformational change of the glycoproteins enabling the virus to bind with the cell and fusing their membranes. The contents of the virus are released into the cytoplasm after which uncoating of the matrix and inner capsid occurs. (2) The single-stranded vRNA is converted to double-stranded DNA using the virus reverse transcriptase enzyme and host cell nucleotides. During reverse transcription the 3' LTR is copied to the 5' end of the viral genome providing identical flanking sequences. (3) The double stranded viral DNA and other associated viral proteins combine to form a pre-integration complex (PIC) necessary for nuclear import. The PIC consists of the reverse transcribed viral cDNA complexed with integrase, matrix, reverse transcriptase, and nucleocapsid proteins (Bukrinsky et al., 1993; Farnet & Haseltine, 1991b; Fassati & Goff, 2001; M. D. Miller, Farnet, & Bushman, 1997). (4) The viral integrase enzyme performs end-processing of the viral DNA, cleaving two nucleotides leaving a CA dinucleotide sticky-end and binds to attachment sites at the terminal ends of the viral genome. (5) Following nuclear import, integrase then makes a staggered cut in the host-cell genome and facilitates strand-transfer and insertion of the viral DNA. (6) The integrated viral genome or provirus is then transcribed by host-cell machinery to produce viral mRNA which is exported from the nucleus. (7) The viral mRNA is then translated by host-cell machinery to produce the components necessary for virus production. (8) The virus components along with two copies of the single-stranded vRNA genome are the transported to the cells plasma membrane where they

assemble to form a new virus. (9) An immature virion then buds off of the cells membrane which then undergoes additional processing by the protease enzyme to produce an infectious mature virion.

The components necessary to support the viral life-cycle are encoded in the viral genome. As illustrated in Figure 3A, the genome of the HIV-1 virus is flanked by two long terminal repeats (LTRs) each containing an untranslated 5' and 3' segment (U5/U3) flanking a repeat region (R) which are necessary for integrase mediated insertion into a host genome. The U3 segment contains the promoter elements necessary for transcription of the entire proviral genome with preference for initiation at the 5' LTR and the U5 segment providing the polyadenylation signal necessary for termination and cleavage at the 3' LTR (Klaver & Berkhout, 1994). The terminal ends of the 5' and 3' LTRs in their U3 and U5 segments, respectively, contain DNA attachment sites necessary for integrase binding and end-processing. Downstream of the 5' LTR is a psi or packaging sequence (Ψ), which is critical for trafficking of the viral genome to the cell membrane for incorporation into viral particles. The backbone of the viral genome codes for the proteins necessary for virus formation, replication and infection with three main genes gag, pol and env. The gag gene codes for the proteins necessary for the viral structure including the inner matrix, capsid, nucleocapsid and other associated proteins (Freed, 1998). The pol gene codes for the enzymes necessary for replication including reverse transcriptase, integrase and protease. The env gene codes for the surface glycoproteins necessary for entry to target cells. The native HIV-1 glycoproteins mediate specificity for CD4+ cells including T lymphocytes, monocytes, dendritic cells and brain microglia

(Wyatt & Sodroski, 1998). In addition to the three main genes, there are two regulatory genes, rev and tat. The rev gene codes for the regulator of virion expression protein (Rev) which is necessary for export of the viral mRNA from the nucleus via binding to a Rev response element (RRE) (Fischer, Huber, Boelens, Mattaj, & Luhrmann, 1995). The tat gene codes for the trans-activator of transcription (tat) protein which acts to increase provirus transcription necessary for viral replication (Dull et al., 1998). Lastly, there are four accessory genes, vif, vpr, vpu, and nef, whose roles are to suppress cellular-based restriction and immunity in order to maintain an active infection (Malim & Emerman, 2008).

LV were designed from HIV-1 retaining the ability to enter a target cell, process the viral genome, facilitate transport to the nucleus, integrate into the host cell's DNA, and have their transgene of interest transcribed while eliminating their ability to replicate which could lead to an active infection. In order to make LV clinically applicable, non-essential sequences and genomic regions involved with viral replication and virulence were removed from the genome including the accessory proteins, nef, vif, vpr, and vpu (Figure 3A). The tat gene was also removed and compensated for by placing a constitutively active promoter at the 5' end of the vector in lieu of the 5' LTRs U3 region for vector production (Dull et al., 1998). The remaining components necessary for vector particle production are placed in trans to each other without the packaging signal to reduce the risk of recombination and production of a replication competent vector. The vector back bone with the flanking LTRs is replaced with a transgene of

interest and any necessary promoter or enhancer elements along with the packaging signal (Figure 3b).

Vector particles are generated using a series of plasmids that express the vector genome and the viral proteins required for particle formation. The technology for packaging vector particles continues to evolve, but the commonly used “third generation” systems utilize a series of four vector plasmids that are introduced into cells by transient transfection (Figure 3B) (Dull et al., 1998). In addition to the transgene plasmid containing the vector backbone or genome, the system uses a plasmid expressing the gag and pol gene regions that produce the HIV-1 structural proteins required for capsid formation and genome integration.

A plasmid expressing HIV-1 rev is also included to activate the rev responsive element engineered into the transgene and gag/pol plasmids. This facilitates nuclear transport and is also included as a safety feature by making it only available during production for expression of packaging components (Dull et al., 1998).

The fourth plasmid expresses the env gene coding for the envelope glycoprotein that engages receptors on the target cells. As the native HIV-1 glycoprotein is generally restricted to CD4 positive cells, investigators utilize alternative envelopes, most commonly the Vesicular Stomatitis Virus G glycoprotein (VSV-G), to facilitate uptake into a wide variety of species and cell types (Aiken, 1997).

Figure 2. Life cycle of a wild-type lentivirus. A brief overview of the major steps involved with infection and replication in the life cycle of a wild-type lentivirus as described in the text.

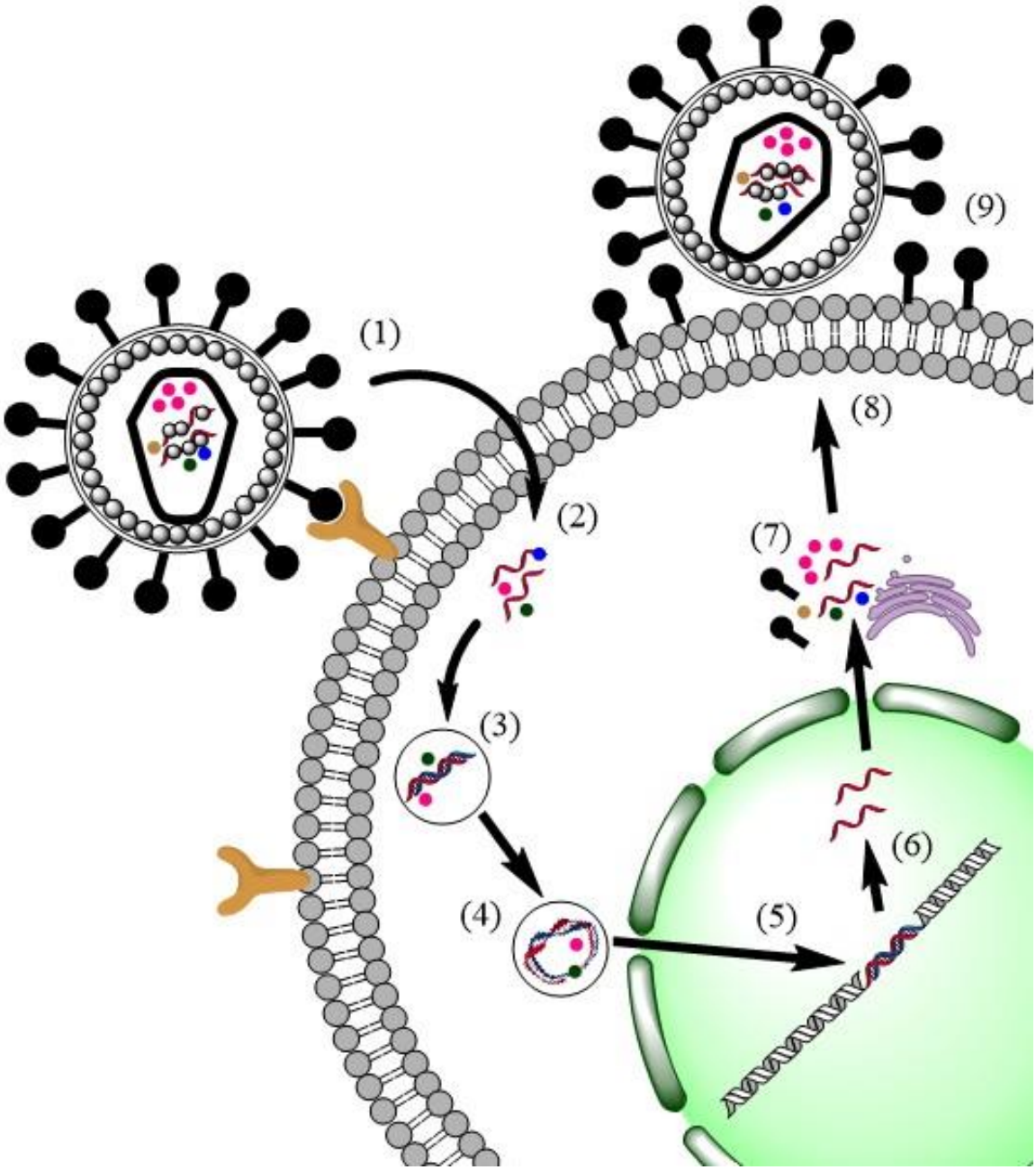
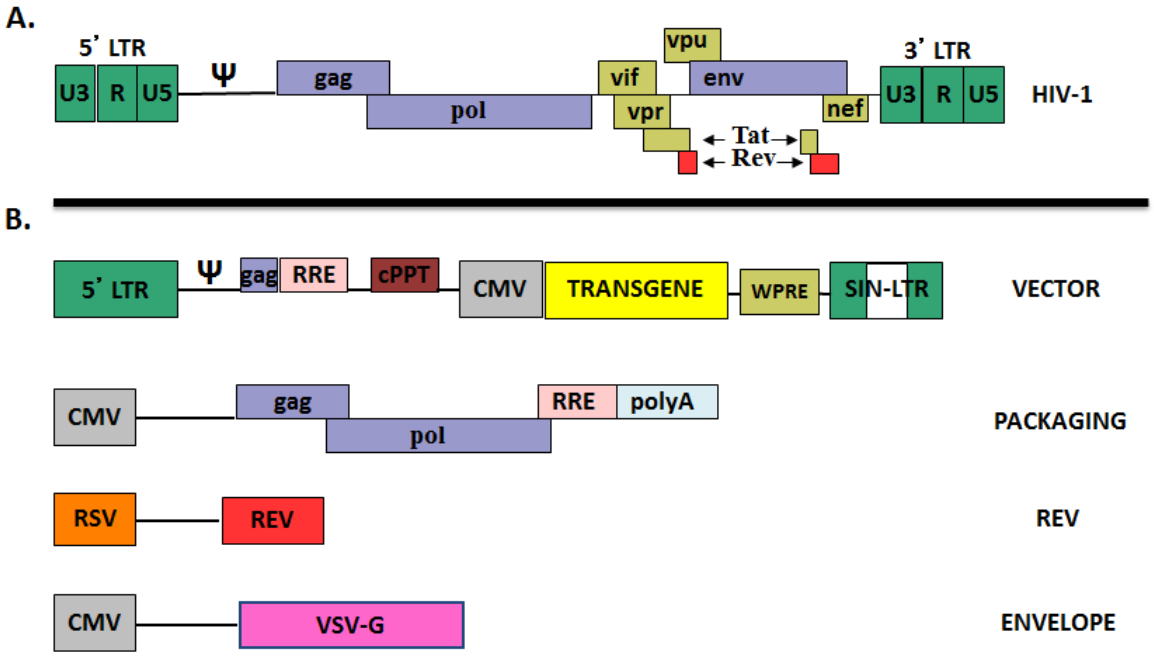


Figure 3. Schematic of HIV-1 and Third Generation Lentiviral Packaging System. (A) The HIV-1 Virus contains three gene regions gag, pol, and env along with accessory proteins and the flanking LTR; (B) The lentiviral components found in the four plasmids used in generating third generation lentiviral vectors. The vector plasmid contains a self-inactivating 3' LTR (SIN-LTR), a Rev responsive element (RRE), a central polypurine tract (cPPT), and the Woodchuck Hepatitis Virus Post-Transcriptional Response Element, a psi sequence (Ψ) for efficient incorporation of the vector RNA genome into particles, and a promoter demonstrated here by the CMV early promoter for transgene expression. The packaging plasmid expresses the gag and pol gene regions of HIV-1 which encode proteins required for virion formation and vector processing. A plasmid expressing rev is provided to facilitate nuclear transport of RRE containing transcripts. The fourth plasmid is the envelope plasmid. LV are commonly pseudotyped to increase the range of cell types and animal species susceptible to vector transduction.



The use of multiple plasmids and the requirement for rev are included to minimize recombination events that would lead to the development of a replication competent virus. For efficient integration, viral particles must contain the proteins encoded in the HIV-1 pol region which are necessary for vector processing including reverse transcriptase, integrase and protease. A short portion of gag sequence which is critical for generating high titer vector, by contributing to the structure of the packaging signal, is also retained in the transgene plasmid (Dull et al., 1998; McBride & Panganiban, 1996). The transgene plasmid contains the minimal components of HIV-1 required for vector production and integration. The transgene of interest is 3' to a promoter element that regulates expression. Additional elements are added to increase vector production and/or expression including a polypurine tract (PPT) (Barry et al., 2001; Van Maele, De Rijck, De Clercq, & Debyser, 2003) and the Woodchuck Hepatitis Virus Post-Transcriptional Response Element (WPRE) (Dupuy et al., 2005; Zufferey, Donello, Trono, & Hope, 1999).

An important safety feature of most LVs is the inclusion of a Self-Inactivating Long Terminal Repeat (SIN-LTR). This feature minimizes the risk of producing a replication-competent lentivirus by recombination with wild-type viruses. The mechanism involves taking advantage of the normal replication cycle of HIV-1. In wild-type HIV-1, the viral promoter is within the U3 region of the 5' LTR and is required to generate the full length viral transcript. The U3 region is also present in the 3' LTR but is not essential in the DNA form of the virus. During viral replication, the RNA genome is reverse transcribed and the 3' LTR is utilized in formation of both the 5' and 3' LTR of the

daughter virus. By incorporating a large deletion including the promoter elements into the U3 region of the 3' LTR any progeny will contain two inactivated LTR after reverse transcription (Miyoshi, Blomer, Takahashi, Gage, & Verma, 1998; Zufferey et al., 1998). Transgene expression is then dependent solely on the internal promoter (for example, the CMV promoter engineered into the vector plasmid as illustrated in Figure 3B).

Another safety element included in the design of lentiviral vectors is the use of insulator elements. Insulator elements reduce the chance of vector induced activation of genes near insertion sites by shielding the transduced genome from enhancers included in vector design and incomplete vector termination reading through into the genome (Ramezani, Hawley, & Hawley, 2008)

LVs are generated by introducing the transgene and packaging plasmids into producer cells, most commonly HEK 293T cells (Figure 4). Vector supernatant is collected from the media and typically contains between 10^5 and 10^7 infectious units per milliliter. This is dependent on the vector design as the addition or deletion of elements to vector design can affect vector titer. After collection vector supernatant is DNase treated to remove any residual plasmid DNA prior to downstream processing (Sastry, Xu, Cooper, Pollok, & Cornetta, 2004). The removal of plasmid DNA is essential to prevent exposing LV treated patients to HIV-1, which could pose the risk of recombination with endogenous retroviral sequences resulting in a replication competent lentiviral vector. LV can be concentrated by ultracentrifugation and clinical vector products are usually purified using a combination of chromatography, tangential

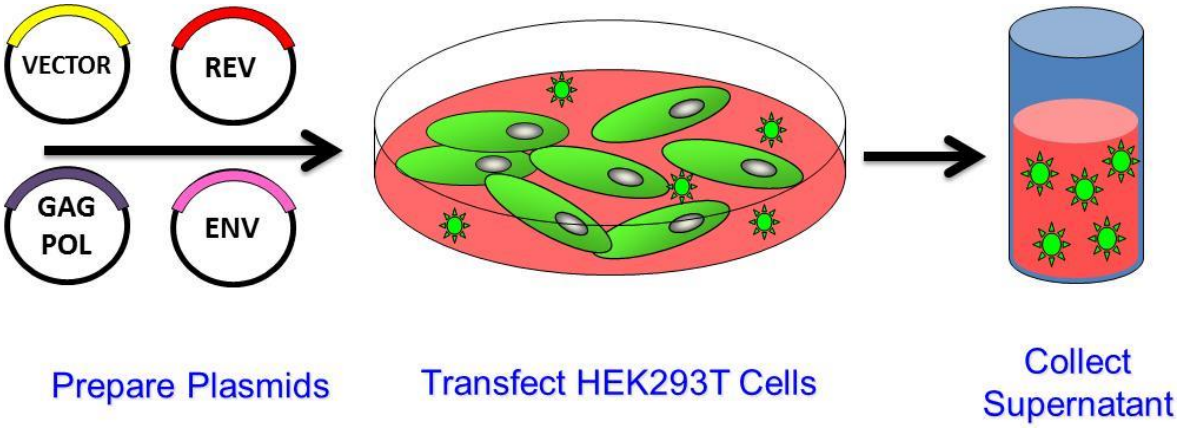
flow filtration and diafiltration (Leath & Cornetta, 2012; Merten et al., 2011; Slepushkin et al., 2003).

In addition to the development of integrating LV, investigators have also sought to develop non-integrating lentiviral vectors (NILV) for applications where integration is not necessary or warranted. Similar to other non-integrating vectors, such as adenoviral vectors and adeno-associated viruses, NILV express their genome episomally as either the linear transcript product from the reverse transcribed vRNA, a 1-LTR circle formed through homologous recombination of the vectors LTRS, or a 2-LTR circle formed through non-homologous end-joining at the terminal ends of the LTRs. In order to effectively transduce a target cell, both integrating LV and NILV must retain the ability to readily enter the cell, form a pre-integration complex, be transported into the nucleus and efficiently express its genetic payload. Depending on the envelope pseudotype used the membrane bound LV particles enter cells either by direct fusion with the plasma membrane (Clapham & McKnight, 2002) or via a receptor-mediated endosomal pathway (Aiken, 1997). In the direct fusion pathway the LV is uncoated upon entry to release the viral contents into the cytoplasm. This allows for reverse transcription of the viral RNA into linear cDNA and development of the pre-integration complex (PIC). The endosomal pathway is dependent upon the pH within the endosome for membrane fusion, subsequent uncoating, and PIC formation within the cytoplasm. The transportation of the PIC to the nucleus is not completely understood, but is believed to occur by an ATP-dependent process (Bukrinsky et al., 1992) via nucleoporins (De Rijck, Vandekerckhove, Christ, & Debyser, 2007; Stevenson, 1996) using nuclear localization signals and cellular

transport mechanisms (Fassati, 2006). Certain of the known localization signals have been removed during LV design; nevertheless, the transduction of quiescent cells by LV is well documented.

Understanding the processing of the LV cDNA after reverse transcription is important when designing NILV. Normally the linear LV cDNA generated during reverse transcription enters the nucleus with the linear 2-LTR form (Chun et al., 1997) representing the preferred substrate for integration (Folger, Wong, Wahl, & Capecchi, 1982; L. Li et al., 2001). A small portion of LV genomes can persist episomally as linear cDNA, 2-LTR circular forms or 1-LTR circular forms. The majority of 2-LTR circles are formed through non-homologous end-joining (NHEJ) of the 5' and 3' LTRs (Jeanson et al., 2002; Kilzer et al., 2003; L. Li et al., 2001). The majority of 1-LTR circles are formed by homologous recombination (HR) between the LTRs (Farnet & Haseltine, 1991a; Gianni, Smotkin, & Weinberg, 1975; Gilboa, Mitra, Goff, & Baltimore, 1979; Jacque & Stevenson, 2006; Kilzer et al., 2003; Shank et al., 1978), as an aberrant product of incomplete reverse transcription (Dina & Benz, 1980; Junghans, Boone, & Skalka, 1982; Kantor et al., 2011; Klarmann, Yu, Chen, Dougherty, & Preston, 1997; Ringold, Yamamoto, Shank, & Varmus, 1977), or through alternative mechanisms such as autointegration (Shoemaker et al., 1980). It is these episomal forms that allow for the stable expression of the vector transgene without integration in non-dividing cells (Apolonia et al., 2007; Bayer et al., 2008; Philippe et al., 2006; Rahim et al., 2009).

Figure 4. Generation of Lentiviral Vector by Transient Transfection. The four packaging plasmids are transfected into cells that have a high capacity for vector production. The most commonly used cell line is HEK 293T. Maximal vector production occurs 48–72 h after transfection. The vector particles are released into the media which is collected and clarified of cell debris. Vector particles can be further purified and/or concentrated.



B. Development of non-integrating lentiviral vectors

"Adapted from Shaw A, Cornetta K. Design and Potential of Non-Integrating Lentiviral Vectors. Biomedicines. 2014; 2(1):14-35"

When designing NILV it is important that modifications maintain the vectors ability to enter their target cells, perform reverse transcription, transport the PIC into the nucleus and efficiently express their transgene product. This entails selectively inhibiting or altering only the aspects of the vectors lifecycle that lead to integration. As integrase-mediated catalysis is the primary means of integration for LV, the inhibition of its function is necessary in the development of NILV.

In wild-type lentiviral vectors integration into the target cell's genome is mediated by the viral integrase. This protein first binds to the viral cDNA at attachment sites located within the U3 region of the 5' LTR and the U5 region of the 3' LTR (Hindmarsh & Leis, 1999; Kulkosky & Skalka, 1994; Masuda, Kuroda, & Harada, 1998; H. Zhou, Rainey, Wong, & Coffin, 2001). Integrase processes the 3' ends of the viral cDNA leaving a CA dinucleotide overhang (Craigie, Fujiwara, & Bushman, 1990; Katz, Merkel, Kulkosky, Leis, & Skalka, 1990) and then attaches the recessed ends to the 5' phosphorylated ends of a double-stranded cut made in the target genome. Integrase then repairs the gaps resulting in a 5 base-pair repeat flanking the inserted vector genome (Katz et al., 1990; Mizuuchi, 1992). There are three general points that can be targeted in developing a NILV; mutations in the integrase protein that alter its ability to

process the target cell chromosomal DNA or alteration in the vectors 5' or 3' LTR attachment sites that prevent integrase from binding and facilitating integration.

The most common target for inhibiting viral integration is mutation of the integrase protein, producing integrase-defective lentiviral vectors. LV integrase is coded for by the HIV-1 Pol region and the region cannot be deleted as it encodes other critical activities including reverse transcription, nuclear import, and viral particle assembly (Gallay, Hope, Chin, & Trono, 1997; Zhu, Dobard, & Chow, 2004). Mutations in *pol* that alter the integrase protein fall into one of two classes: those which selectively affect only integrase activity (Class I); or those that have pleiotropic effects (Class II) (Engelman, 1999). Mutations throughout the N and C terminals and the catalytic core region of the integrase protein generate Class II mutations that affect multiple functions including particle formation and reverse transcription (Engelman, 1999; Saenz et al., 2004; Wiskerchen & Muesing, 1995). Therefore, class II mutations are not suitable when designing NILVs because they disrupt functions that are critical for vector processing and expression.

Class I mutations limit their effect to the catalytic activities, DNA binding, linear episome processing and multimerization of integrase (Banasik & McCray, 2010). The most common Class I mutation sites are a triad of residues at the catalytic core of integrase, including D64, D116, and E152 (Kulkosky, Jones, Katz, Mack, & Skalka, 1992; Shibagaki & Chow, 1997). Each mutation has been shown to efficiently inhibit integration with a frequency of integration up to four logs below that of normal

integrating vectors while maintaining transgene expression of the NILV (Apolonia et al., 2007; Leavitt, Robles, Alesandro, & Varmus, 1996; Nightingale et al., 2006).

Another alternative method for inhibiting integration is mutations in the integrase DNA attachment sites (LTR *att* sites) within a 12 base-pair site of the U3 or an 11 base-pair site of the U5 regions at the terminal ends of the 5' and 3' LTRs, respectively (Apolonia et al., 2007; Brown, Chen, & Engelman, 1999; Masuda et al., 1998; Masuda, Planelles, Krogstad, & Chen, 1995; Nightingale et al., 2006). These sequences include the conserved terminal CA dinucleotide which is exposed following integrase-mediated end-processing. Single or double mutations at the conserved CA/TG dinucleotide result in up to a three to four log reduction in integration frequency (Nightingale et al., 2006); however, the vector retains all other necessary functions for efficient uptake and expression.

While NILV can significantly reduce the frequency of integration there still remains a low level of vector integration (Apolonia et al., 2007; Cornu & Cathomen, 2007; Gaur & Leavitt, 1998; Leavitt et al., 1996; Leavitt, Shiue, & Varmus, 1993; Matrai, Chuah, & VandenDriessche, 2010; Nakajima, Lu, & Engelman, 2001; Nightingale et al., 2006; Philippe et al., 2006; Yanez-Munoz et al., 2006). Integration associated with NILV has been studied using insertion site analysis and high-throughput sequencing. These analyses indicate that the integration observed is not integrase-mediated as the insertion sites lack the canonical features of LTR end-processing including the five base-pair repeat of genomic DNA flanking at the site of vector insertion (Gaur & Leavitt, 1998; Koyama, Sun, Tokunaga, Tatsumi, & Ishizaka, 2013; Matrai et al., 2011; Matrai et al.,

2010; Nightingale et al., 2006). The vector inserts also vary with some containing fully intact sequence, truncations at the terminal ends of the LTRs, or insertions/deletions of genomic DNA flanking the vector. A significant number of these integrations are occurring at sites of chromosomal breakage and are mediated by NHEJ mechanisms (Koyama et al., 2013; D. G. Miller, Petek, & Russell, 2004). It may be possible to impede illegitimate integration of NILV by inhibition of cellular factors in the double-strand break (DSB) repair pathway (Koyama et al., 2013). Whether or not inhibitors to DNA repair can be used clinically to limit non-integrase mediated integration remains to be determined.

Another method for further reducing the frequency of illegitimate integration is limiting the linear form of the vector DNA. Linear DNA has been shown to integrate much more efficiently than supercoiled DNA associated with circular DNA (Folger et al., 1982; L. Li et al., 2001). The linear form also appears to be the preferred substrate for both integrase and non-integrase-mediated insertion. One approach has been to limit linear 2-LTR episomal forms by inducing the formation of circular episomal forms. For example, Kantor and colleagues have shown that deleting the vectors 3' polypurine tract (PPT) results in aberrant reverse transcription leading to the preferential formation of 1-LTR circular episomes and a reduction in linear forms (Kantor et al., 2011). Using this strategy they were able to reduce the frequency of integration by 10-fold when using an integrating vector. Of relevance to NILV, this modification reduced the frequency of integration of an integrase deficient LV by another 3-fold over integrase deficient LV without the modification. As newer modifications are developed to reduce integration

they will need to be tested experimentally to ensure there is no reduction in the level of transgene expression.

It should be noted that LTR att site mutations have been directly compared to point mutations in integrase. The consensus is that mutations to integrase provide a greater reduction in the frequency of integration. Yanez-Munoz et al. estimated the frequency of reversion mutations in NILV to reach 1/815 (Yanez-Munoz et al., 2006). The point mutations to LTR att sites could carry a higher rate of reversions but whether LTR att site mutations with larger deletions will reduce the frequency of illegitimate integration remains to be determined. Interestingly, while mutations to integrase and the LTR att sites independently inhibit integration efficiently, some studies suggest there are no synergistic effects upon combining these mutations to further reduce integration by a vector (Apolonia et al., 2007; Nightingale et al., 2006). However, whether this is due to the type or efficiency of the mutations being combined remains to be determined.

While NILV can effectively inhibit integration, they are associated with a significantly reduced level of transgene expression as compared to a normal integrating LV (Apolonia et al., 2007; Bayer et al., 2008; Cornu & Cathomen, 2007; Coutant, Frenkiel, Despres, & Charneau, 2008; Kantor et al., 2011; Nakajima et al., 2001; Naldini et al., 1996; Philippe et al., 2006; Vargas, Gusella, Najfeld, Klotman, & Cara, 2004; Vargas, Klotman, & Cara, 2008). This remains a key issue in developing clinically effective NILV. One approach for improvement is removing or reducing inhibitors to episomal transgene expression. Bayer et al. have shown that removal of cis-acting sequences within the U3 region of the vectors LTR improves episomal transgene expression by

nearly 3-fold (Bayer et al., 2008). However, other mechanisms of episomal inhibition may be involved because expression was still below that of the normal, integrating LV control.

Other approaches include the inhibition of cellular restriction factors (Berger, Goujon, Darlix, & Cimarelli, 2009; Negri et al., 2012), codon optimization to improve protein production and potency (Negri et al., 2007; Suwanmanee et al., 2013), and the use of histone deacetylase inhibitors for transgene activation (Pelascini, Janssen, & Goncalves, 2013).

Combining modifications that increase transgene expression and reduce integration will be needed to maximize the safety profile of NILV. If expression is low, a higher number of NILV will be required per cell to obtain the therapeutic benefit. The higher numbers of vector episomes will in turn increase the chance of illegitimate integration. For certain gene therapy trials, such as those using ex vivo gene transfer of CD34+ hematopoietic cells, the number of cells treated may exceed 5×10^8 cells for an adult. Even with a four log reduction in integration, a significant number of cells will contain integrated proviruses. Therefore efforts to minimize integration and optimizing expression should be considered for both therapeutic and safety reasons.

C. Insertion site analysis

A hallmark of LV is their ability to stably integrate their genome into a host cell's DNA. Integrase mediated insertion by LV while theoretically random in nature, demonstrates preferential insertion in gene coding regions (Schroder et al., 2002). This presents the risk of insertional mutagenesis which could lead to aberrant gene expression and possible oncogenesis. Site-directed integration systems can offer a safer alternative to integrase-mediated insertion, though they have demonstrated some off-target insertions which must be accounted for to improve their clinical utility (Beumer, Bhattacharyya, Bibikova, Trautman, & Carroll, 2006; J. C. Miller et al., 2007). Several methods have been employed for the analysis of insertion sites of wild-type LV, however NILV offer more difficulty in their capture and analysis. Whereas wild-type LV insert by integrase-mediated integration with intact vector sequences and canonical features including end-processing and a five base-pair flanking repeat, NILV insertions are much more variable. NILV have been shown to integrate by non-integrase-mediated mechanisms such as HR and/or NHEJ (Gaur & Leavitt, 1998; Matrai et al., 2011; Nightingale et al., 2006). NILV insertion sites present insertions and/or deletions at the vector-genome junction which could inhibit many of the PCR-based methods for vector insertion site analysis, which are dependent on intact primer sites located at or near these junctions. In order to develop vectors with safer insertion site profiles and to evaluate the integrity of insertion sites post-transduction, it is necessary to have

methods in place which can evaluate them with a high degree of sensitivity and in a high-throughput manner while accounting for variables in the vector-genome junction.

Previous techniques for the capture of vector integration sites were either very labor intensive and/or lacked the sensitivity necessary to evaluate the large number of insertion sites possible in many applications. One technique which has been applied that can effectively evaluate the variable insertion sites of NILV employs the use of inverse PCR followed by insertion of the PCR product into cloning vectors (Nightingale et al., 2006). However this type of technique is very labor intensive and limits the number of insertion sites which can be analyzed in a high throughput manner. Another method that has been utilized for capturing vector insertion sites is ligation-mediated PCR (LM-PCR) (Mueller & Wold, 1989). However this method lacks the sensitivity to capture a large number of integrations or characterize the variability associated with NILV insertion sites. Two of the most sensitive methods for analyzing vector insertion sites are linear amplification-mediated PCR (LAM-PCR) (Schmidt et al., 2007) and nonrestrictive LAM-PCR (nrLAM-PCR) (Gabriel et al., 2009). These methodologies have been combined with next-generation sequencing to successfully analyze large numbers of insertion sites in a relatively high-throughput manner. However, these methods have their limitations as well. LAM-PCR utilizes restriction enzymes and is therefore dependent on the presence of restriction sites in proximity to the vector insertion site and can result in missing up to 50% of insertion sites (Gabriel et al., 2009). A major hurdle with nrLAM-PCR is the necessity for a single stranded ligation of a linker cassette. The technology available for this step is very inefficient as most enzymes with this

capability are intended for use with single-stranded RNA templates. An important obstacle with both of these methods for the elucidation of NILV insertion sites is the necessity for LTR-specific primers to amplify the vector-genome junction. Deletions at the vector-genome junction are common with NILV which could result in the loss of retrieving these insertion sites depending on the size of the deletion. As new technologies have emerged, improved methodologies have been developed for analyzing vector insertion sites that can bypass many of these limitations. These techniques utilize sonication to shear the vector transduced DNA to fragments in a size range which allows for easier manipulation and analysis. Some methods which utilize this technique still rely on the use of LTR-specific primers for the capture, purification and analysis of insertion sites (De Ravin et al., 2014; S. Zhou et al., 2014). However, utilizing on optimal sheared fragment size and modifying the primer site to sit farther back within the vector can allow for capture of a larger percentage of NILV insertion sites which contain deletions at the vector-genome junction. As high-throughput next generation sequencing technologies advance and allow for the processing of larger templates the analysis of insertion sites without the necessity to depend on primer or restriction sites will allow a more accurate evaluation which can improve the safety of these vectors for use in the clinic.

D. Statement of Purpose

The inherent risks associated with using viral vectors in patients along with adverse side effects observed with early retroviral vector trials necessitates the continuous improvement of lentiviral vector mediated gene therapy to facilitate its safety for use in the clinic. We have evaluated modifications in vector production to improve the purity of the vector supernatant, designed novel non-integrating lentiviral vectors with multiple modifications to vector design, and optimized methodologies for analyzing vector insertion sites post-transduction. We hypothesized that by utilizing the results from these studies we can improve the safety of lentiviral vector mediated gene therapy by incorporating more clinically relevant reagents into vector production, developing NILV with reduced frequencies of illegitimate integration, and improving our ability to analyze the aberrant insertion sites associated with NILV. This was accomplished through the following specific aims:

1. <u>Improving the purity of lentiviral vector supernatant through incorporation of an FDA approved deoxyribonuclease (DNase) into vector production.</u>

This aim will test the hypothesis that incorporating an FDA approved recombinant human DNase into lentiviral vector production for removing residual plasmid DNA will improve the purity and clinical potential of the final vector supernatant product. By applying quantitative PCR (Q-PCR) we will determine whether a

DNase, FDA approved for use in treating patients with cystic fibrosis, will provide a more clinically relevant alternative to a previously described bacterial-derived DNase.

2. Combining multiple modifications to NILV design to improve their safety and efficacy for clinical applications.

By using flow cytometry, an antibiotic resistant colony formation titer assay, and Q-PCR copy number analysis, the aim will demonstrate that novel NILV combining multiple modifications for inhibiting vector integration and a deletion mutation for improving transgene expression results in NILV backbones with safer profiles for use in applications where integration is not necessary.

3. Modifying vector insertion site capture methodologies to improve the analysis of aberrant NILV insertion sites.

By utilizing optimized vector insertion site capture methodologies coupled with high-throughput next-generation sequencing, this aim will improve the ability to analyze the integration profiles of vectors while taking into account the aberrant vector-genome junctions associated with NILV.

The following chapters within this dissertation describe the methods and results of these specific aims.

II. Using Pulmozyme DNase treatment in lentiviral vector production

"Reprinted with permission from HUMAN GENE THERAPY: METHODS 2012, published by Mary Ann Liebert, Inc., New Rochelle, NY"

A. Introduction

Lentiviral vectors have shown promise as a new therapeutic tool for patients with life threatening disorders (Cartier et al., 2009; Cavazzana-Calvo et al., 2010; DiGiusto et al., 2010; Kalos et al., 2011; Levine et al., 2006). Most investigators generating lentiviral vectors utilize a transient transfection when manufacturing these vectors for research and clinical applications (Dull et al., 1998; Gasmi et al., 1999; X. Lu et al., 2004). Vector components are separated into different plasmids to reduce the risk of recombination and the potential production of a replication competent lentivirus (RCL) (Dull et al., 1998; Zufferey et al., 1998), a strategy that to date appears effective (Cornetta et al., 2011). A concern with the transient transfection method is the residual plasmid DNA remaining in the final product, potentially exposing patients to plasmids expressing HIV-1 DNA (Zufferey, 2002). To decrease this risk, the DNase Benzonase has been shown to significantly decrease plasmid DNA in lentiviral vector preparations (Sastry et al., 2004). Unfortunately, Benzonase is a bacterial derived product and does not have FDA approval, a factor that will be problematic as vectors move from early phase studies into licensed products. In contrast, Pulmozyme is a recombinant human DNase (rhDNase) that is FDA approved for the treatment of cystic fibrosis (Pan,

Sinicropi, & Lazarus, 2001; Shak, Capon, Hellmiss, Marsters, & Baker, 1990; Shire, 1996; Ulmer et al., 1996) and could serve as a safer alternative.

The timing of DNase treatment is also an area for potential improvement. DNase is generally added to the vector product after the final harvest, increasing processing time that can decrease the potency of the final product. Moreover, the DNase will remain in the final product unless additional processing is performed. We therefore compared DNase treatment of the final product versus adding DNase earlier in the production process as a potential means to minimize residual DNase in the final product. Our findings demonstrate that Pulmozyme is an effective DNase and a suitable alternative to Benzonase for removing plasmid DNA from lentiviral vector products.

B. Methods

Cell culture and reagents

HEK293T cell lines were used for vector production and HEK293 cells were used for transduction and subsequent infectious titer measurement. Both cell lines were obtained from the American Type Culture Collection (ATCC, Manassas, VA) and were maintained in D10 medium [88% high-glucose Dulbecco's modified Eagle's medium (DMEM, Invitrogen, Carlsbad, CA), 10% fetal bovine serum (FBS, HyClone, Logan UT), 100 units/ml penicillin-streptomycin (Pen Strep, Invitrogen) and 2 mM Glutamax (Invitrogen)] incubated at 37°C with 5% CO_2. Viral supernatants were harvested in OptiPRO Serum Free Medium (Invitrogen).

Vector production

Vector was generated in HEK 293T cells plated at 2 X 10^6 / T25 flask in D10 medium. After approximately 24 hours the medium was changed to fresh D10 with the subsequent addition of the transfection mixture using the Promega Profection mammalian transfection kit (Promega, Madison, WI) according to the manufacturer's instructions. The transgene vector plasmid used was pcDNA-CS-CGW (4.4 µg/T25 flask) provided by Philip Zoltick (Children's Hospital, Philadelphia, PA), containing the enhanced green fluorescent protein (eGFP) gene driven by the cytomegalovirus (CMV) promoter. The three packaging plasmids used were pMDL expressing gag-pol (2.2 µg/T25 flask), pMDG expressing the vesicular stomatitis virus (VSV)-G envelope

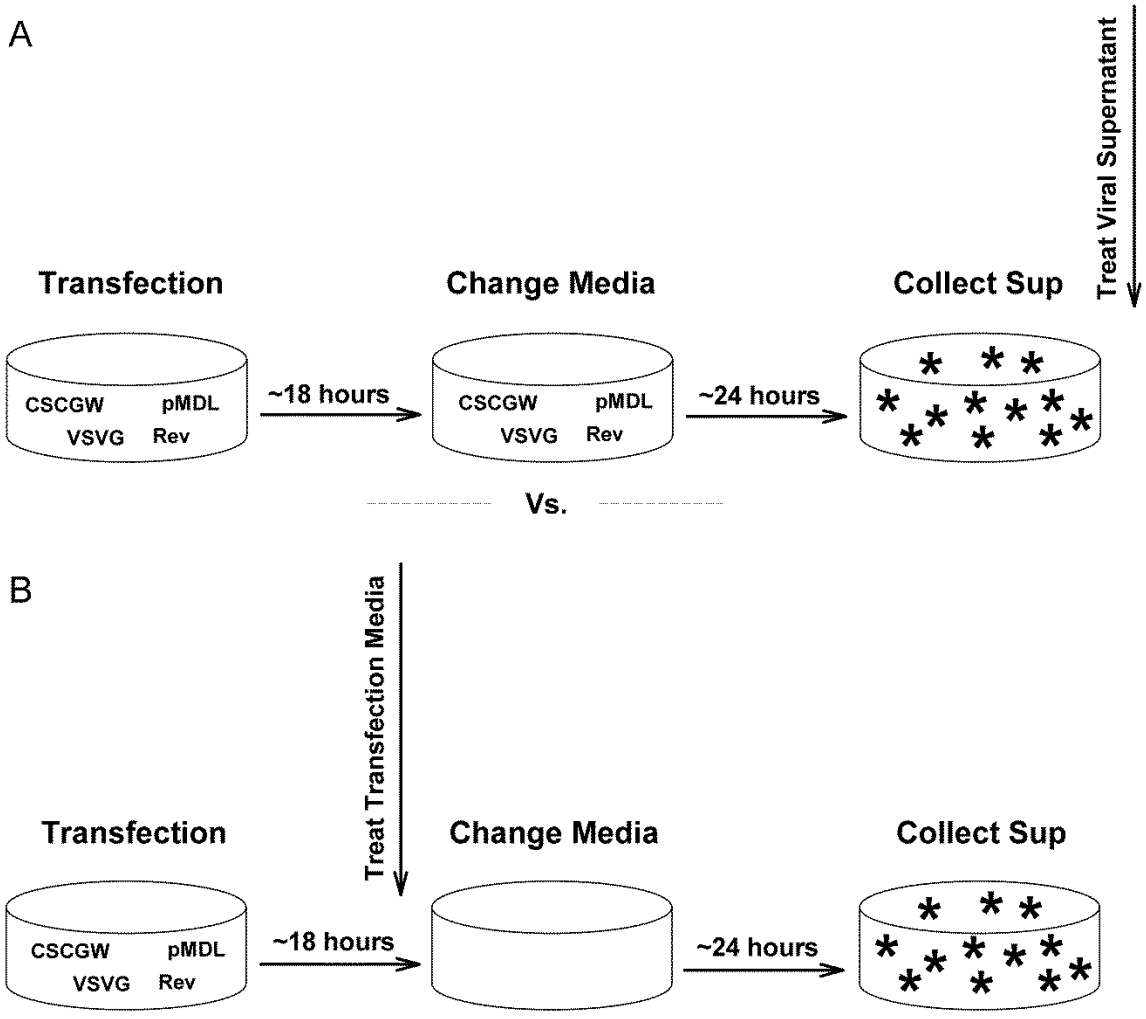
glycoprotein (1.5 μg/T25 flask) and pRSV-REV for Rev expression (1.1 μg/T25 flask) (Dull et al., 1998; Zufferey et al., 1998). The cells were then incubated at 37° C for 16 – 18 hours after which time the transfection medium was replaced with OptiPRO serum-free medium. Twenty four hours later the viral supernatant was collected and clarified using a 0.45 μm filter and stored at -80° C.

DNase treatment

Enzymes used for plasmid DNA removal were Benzonase nuclease (Novagen, Darmstadt Germany), and pharmaceutical grade Pulmozyme (Dornase alpha; Genentech, Inc., San Francisco, CA). DNase treatment was tested at two time points (Figure 5). In Protocol A DNase is added to the final product after vector harvest and clarification. In Protocol B, DNase treatment was added at the end of an 18 hour transfection and immediately prior to the final medium change. Vector was incubated with the DNase at 37° C with 5% CO_2. The incubation times and DNase concentrations for both protocols are described in Results.

Figure 5. Protocols for DNase treatment during lentiviral vector production.

(A) Protocol A illustrates that DNase is added to harvested vector at the end of production. (B) Protocol B illustrates the addition of DNase at the end of the transfection period and prior to medium change and subsequent harvest.



Residual DNA purification

DNA was extracted by a phenol/chloroform purification followed by ethanol precipitation in order to completely deactivate any residual DNase. An equal volume of phenol:choloroform:isoamyl alcohol (25:24:1; Sigma-Aldrich, St. Louis, MO) was added to the viral supernatant. The mixture was then vortexed and spun for 10 minutes at 13,000 RPM. The aqueous phase was removed and an equal volume of chloroform:isoamyl alcohol 24:1 (Sigma-Aldrich, St. Louis, MO) was added. The tube was briefly vortexed and spun for 10 minutes at 13,000 rpm. The aqueous phase was removed and adjusted to 0.3M sodium acetate. Twice the volume of ice-cold 100% ethanol was added and incubated at -80° C for two hours followed by to centrifugation at 4° C for 30 minutes. The pellet was washed with 70% ethanol, centrifuged at 4° C for 30 minutes, air dried, then resuspended in water at 65° C for one hour. The DNA concentration was obtained using the NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE) or with a BioPhotometer (Eppendorf, Hamburg, Germany).

Real Time PCR

Quantitative-polymerase chain reaction (Q-PCR) was performed with the 7500 Real- Time PCR System and analyzed using its 7500 system SDS software (Applied Biosystems, Life Technologies Corporation, Carlsbad, CA). Q-PCR was performed with forward and reverse primers (VSVG-F1: 5'-tgcaaggaaagcattgaacaa-3' and VSVG-R1: 5'-gaggagtcacctggacaatcact-3') and a probe specific for a 120 bp fragment of the VSV-G

envelope sequence (TP-VSVG: 5'-6FAM-aggaacttggctgaatccaggcttcc-TAMRA-3'). Purified plasmid DNA was normalized based on the volume of the supernatant tested by diluting equally to within an optimal range for working concentrations. 2 µl of the normalized plasmid DNA was added to a PCR master mix consisting of 1X Taqman Buffer A, 0.4 mM dNTPs, 1 mM MgCl2, 1.2 µM VSVG-F1 & VSVG-R1 primers, 0.4 µM TP-VSVG probe and 0.025 U/µl AmpliTaq Gold. Reactions were performed in duplicate or triplicate using one cycle of 50° C for 2 min., 95° C for 10 min. and 40 cycles of 95° C for 15 sec. and 60° C for 1 min. A standard curve was prepared on the basis of serial dilutions of a VSVG standard ranging from 10^1 to 10^5 copies.

Measurement of infectious titer

Determination of infectious titer was performed in order to determine the effect of DNase treatment on vector titer. HEK293 cells were transduced in duplicate by plating cells at 10^5 cells per well in six well tissue culture treated plates and incubated overnight at 37°C with 5% CO_2. The next day the medium was removed and a transduction mix (D10 culture medium, 8 µg/ml of polybrene (Sigma-Aldrich), and serial dilutions of 1:100 and 1:1000 of viral supernatant) was added and then returned to the incubator for four hours. The transduction mix was then removed and two ml of fresh culture medium was added and the cells were incubated for approximately 72 hours. Cells were trypsinized, washed in Dulbecco's Phosphate Buffered Saline (DPBS, Invitrogen), pelleted and resuspended in 250 µl of 1% paraformaldehyde in DPBS. In order to calculate infectious titer, GFP expression was measured using a FACScan

(Becton Dickinson Immunocytometry Systems, San Jose, CA) as previously described (Sastry, Johnson, Hobson, Smucker, & Cornetta, 2002).

Analyzing plasmid DNA degradation in serum

DNA for the RD114 envelope plasmid was utilized to analyze the effects of serum and calcium phosphate on DNA degradation. RD114 Plasmid (50 ng/µl) was incubated at 37°C with 5% CO_2 in serum free DMEM medium, D10, or complexed with the calcium phosphate mix as described above added to D10 medium. Various incubation times up to four hours were observed. After incubation, 5 mM EDTA was added to inhibit any DNase activity and the results were compared by running the products on an agarose gel.

C. Results

The ability of Pulmozyme to decrease plasmid DNA in lentiviral supernatant was assessed by generating third generation lentiviral vectors using a four plasmid system, including plasmids expressing gag/pol, Rev, the VSV-G envelope, and a transgene plasmid expressing GFP. Vector was generated using calcium phosphate transfection, and DNase treatment of the final product was compared to DNase treatment at the end of the transfection (Figure 5). Plasmid removal was assessed by Q-PCR with primers specific for the VSV-G envelope gene (the VSV-G plasmid represents approximately 1/8th of the total transfected DNA). As a control, 50 U/ml of Benzonase was utilized, as this concentration of DNase has previously been shown to efficiently remove plasmid DNA from lentiviral products (Sastry et al., 2004).

DNA Degradation after Treatment with Pulmozyme®

The degradation of plasmid DNA after vector harvest (Protocol A, Figure 5) was evaluated using a variety of Pulmozyme concentrations ranging from 1 to 50 U/ml. Benzonase, at 50 U/ml, was used as a control. Incubation time was one hour. As shown in Figure 6, concentrations of Pulmozyme at or above 20 U/ml decreased plasmid DNA to within the limits of detection, and plasmid DNA removal by Pulmozyme was equal to or better than the removal obtained with Benzonase. Increasing the incubation time to two or four hours did not improve DNA degradation when treating by this method (data not shown).

Next, the efficiency of DNA degradation was measured when DNase was added at the end of the transfection period and prior to the final medium change (Protocol B, Figure 5). As shown in Figure 7A, Pulmozyme at a concentration of 25 U/ml provided a greater than two log decrease in plasmid DNA and was comparable to or better than 50 U/ml of Benzonase. Increasing the Pulmozyme concentration above 25 U/ml did not provided any added benefit (Figure 7A). Concentrations less than 25 U/ml did not attain a two log reduction (data not shown). Any potential adverse effect of Pulmozyme on vector titer was also assessed. As shown in Figure 7B, the vector titer was not significantly affected by Pulmozyme, even at concentrations up to 100 U/ml.

To determine whether longer incubation times following protocol B would improve the effectiveness of Pulmozyme, the incubation time at the end of the transfection period was increased. As shown in Figure 8A, addition of Pulmozyme at 25 U/ml led a similar amount of residual DNA in the final product after 1, 2, and 4 hours of incubation. To evaluate whether longer incubations allowed for a decrease in Pulmozyme concentrations, a four hour incubation was performed with concentrations varying from 5 to 25 U/ml. As shown in Figure 8B, the lowest DNA content was noted at 20 and 25 U/ml. The 4-hr incubation did not alter the vector titer (Figure 8C). When taking the data in aggregate, adding Pulmozyme at 20 U/ml after transfection and incubating for one hour gave optimal DNA degradation, and results were similar to or better than with Benzonase at 50 U/ml. Unfortunately, the level of DNA degradation when treating after transfection (protocol B) was significantly less than that obtained by treating after vector harvest (protocol A).

DNA degradation in the presence of serum

A significant difference between the vectors treated in protocol A versus the vector treated in protocol B is the presence of serum in the medium. Specifically, in protocol A the vector is treated in serum-free medium, whereas the vector in protocol B is treated in serum-containing medium. To evaluate the effectiveness of DNase in the different media we incubated plasmid DNA in serum-free medium (OptiPRO), and in D10 medium (DMEM with 10% fetal bovine serum), for up to 4 hours (37°C, 5% CO_2). Interestingly, free plasmid DNA is stable in serum-free medium (Figure 9A), whereas plasmid in medium containing fetal bovine serum leads to a time-dependent degradation of plasmid DNA (Figure 9B). Complexing plasmid DNA with calcium phosphate provides some protection from serum degradation (Figure 9C). This is consistent with observations that calcium phosphate can form complexes with DNA that are resistant to extracellular DNase activity (Loyter, Scangos, & Ruddle, 1982). However, complexing plasmid DNA with calcium phosphate does not appear to provide the same degree of protection in the presence of Pulmozyme (Figure 9D). These findings suggest that the lower DNA degradation seen with protocol B is not the result of Pulmozyme inhibition by serum.

Figure 6. Plasmid DNA degradation after 1 hour of DNase treatment post-harvest.

Average Q-PCR results from four experiments using primers for the VSV-G plasmid as a measure of residual plasmid DNA after DNase treatments. X-axis values represent the Unit/ml of DNase for Pulmozyme [P], Benzonase [B] or an untreated control [Un]. Values on the Y-axis represent the number of gene copies per 500 µl of viral supernatant. Error bars indicate the standard deviation (SD) of the mean.

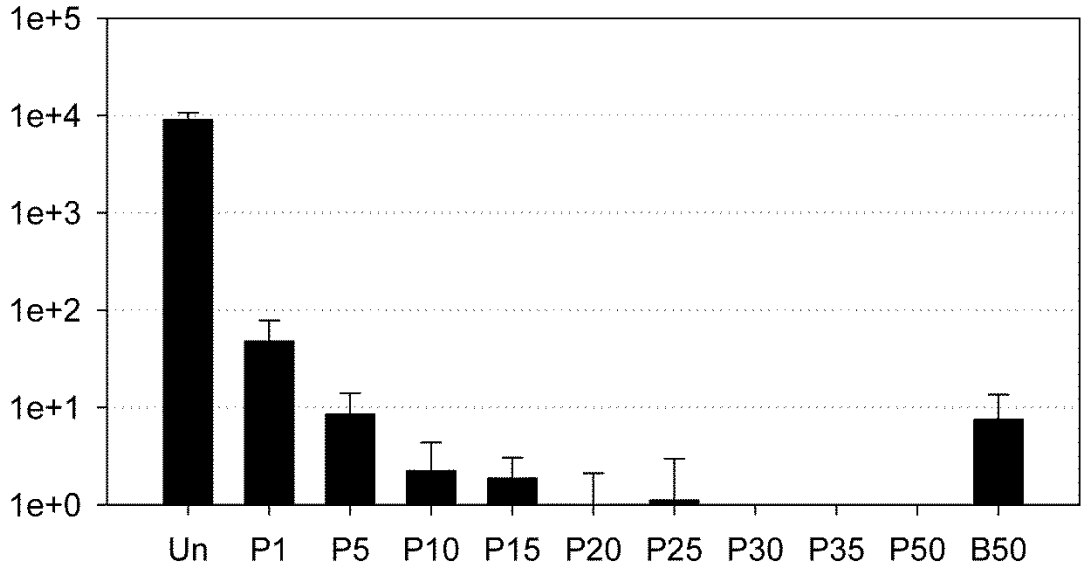


Figure 7. Plasmid DNA degradation after 1 hour DNase treatment performed at the end of transfection. (A) Average Q-PCR results for duplicate experiments using primers for the VSV-G plasmid as a measure of residual plasmid DNA in the final product. DNase treatment was performed at the end of transfection. X-axis values represent the Unit/ml of DNase for Pulmozyme [P], Benzonase [B] or an untreated control [Un]. Values on the Y-axis represent the number of gene copies per 500 µl of viral supernatant. (B) Infectious titer of final vector product as assessed by GFP expression, given as infectious units per ml. Error bars represent SD of the mean.

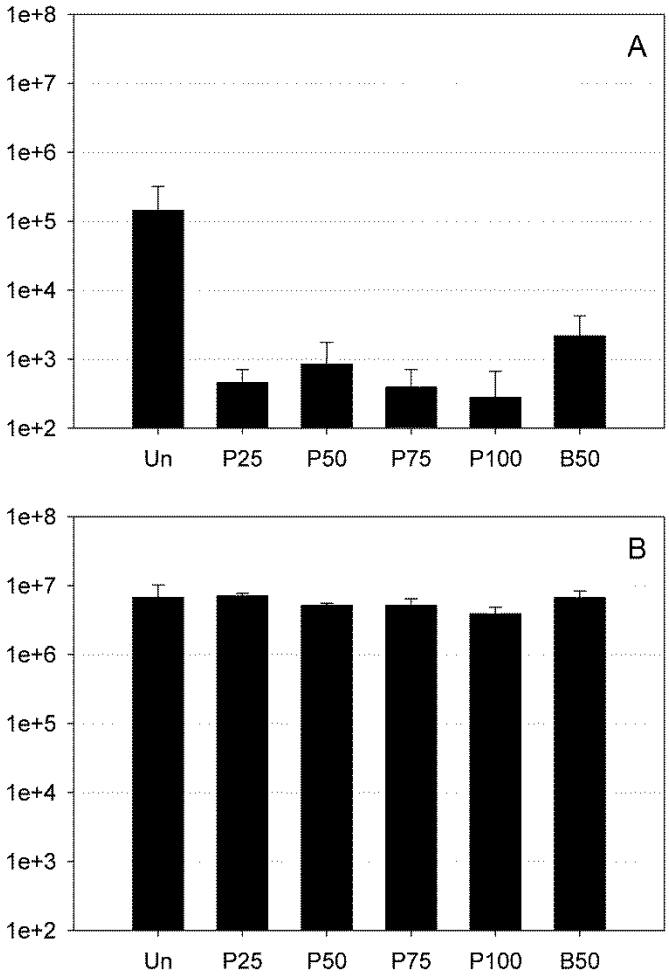


Figure 8. Plasmid DNA degradation over time when treating at the end of transfection. (A) Q-PCR results for VSV-G plasmid in the final product when Pulmozyme (25 U/ml) is added at the end of transfection for 1-4 hours. Values on the Y-axis represent the number of gene copies per 500 µl of viral supernatant. (B) Q-PCR results of residual VSV-G after a 4 hour incubation with various concentrations of DNase. X-axis values represent the U/ml of DNase for Pulmozyme [P], Benzonase [B] or an untreated control [B]. Values on the Y-axis represent the number of gene copies per 500 µl of viral supernatant. (C) Infectious titer of final vector product as assessed by GFP expression, given as infectious units per ml. Error bars represent SD of the mean.

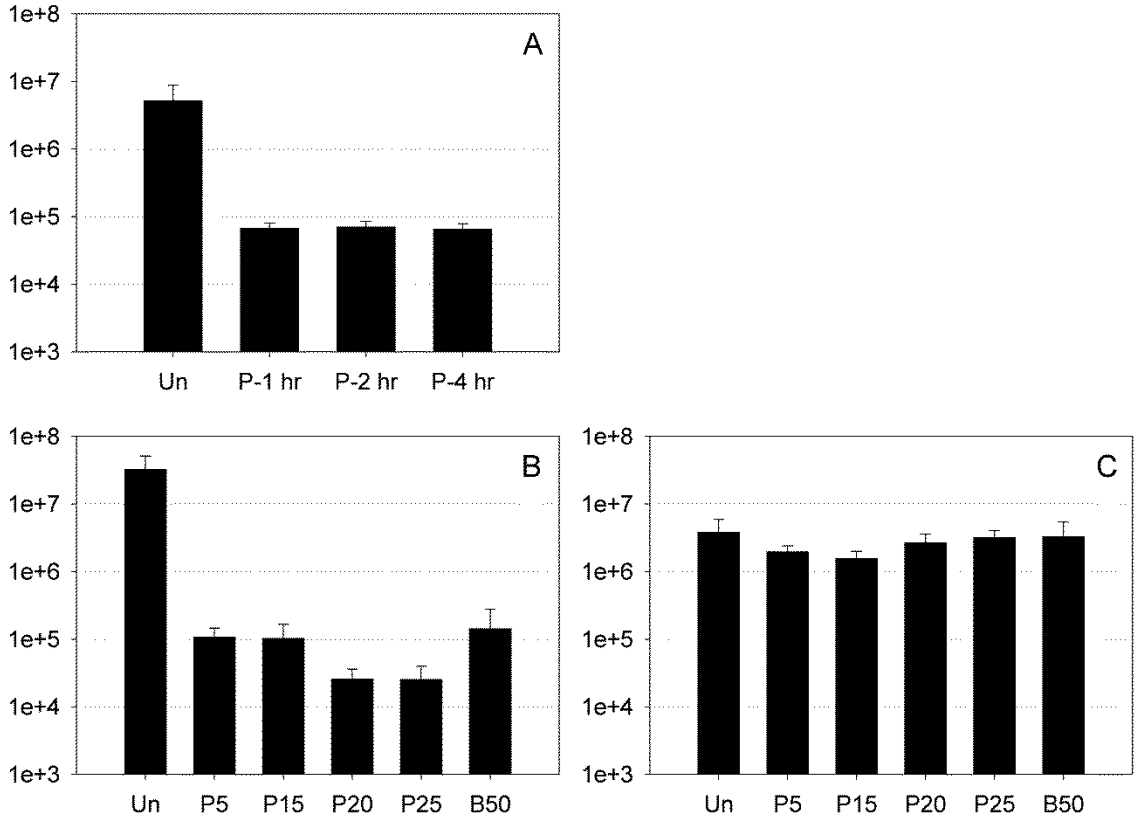
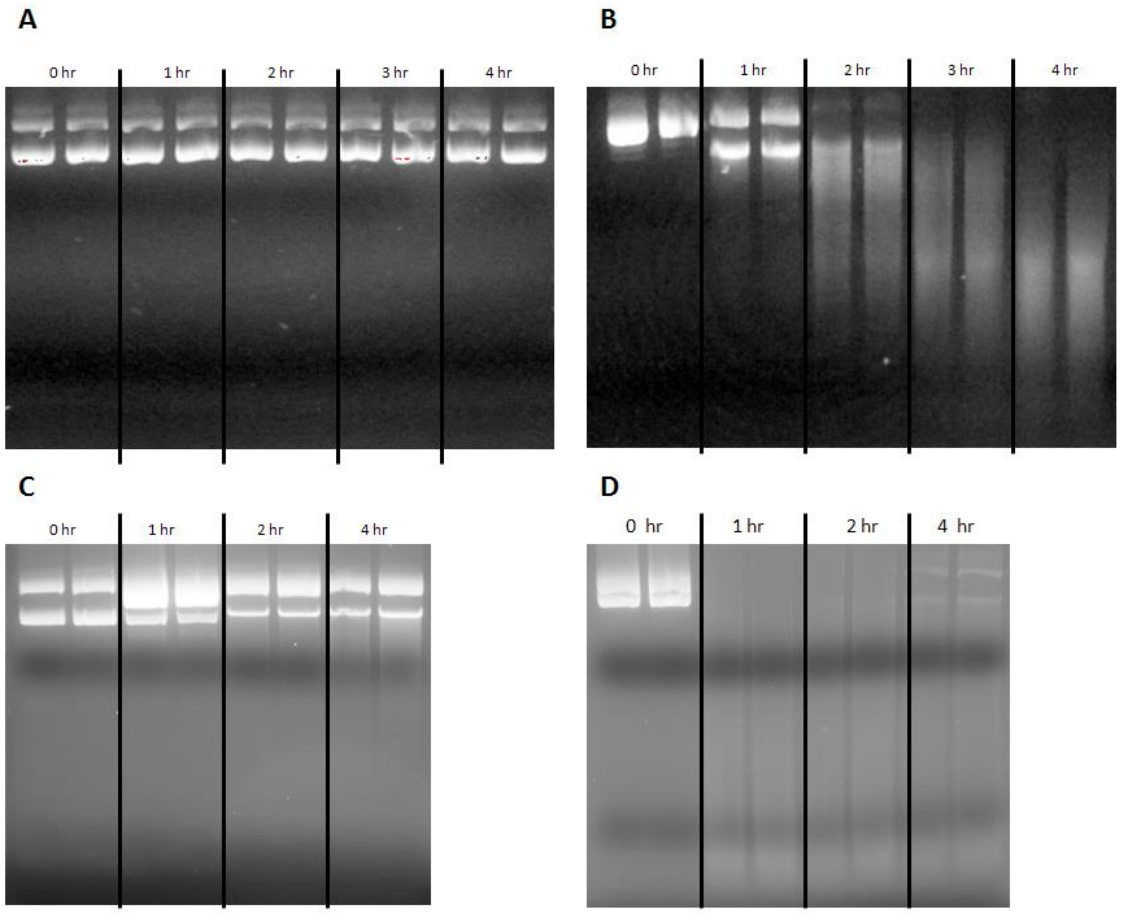


Figure 9. The effects of serum and calcium phosphate on plasmid DNA degradation. Plasmid incubated in (A) serum free medium (OptiPRO); (B) D10 (serum containing) medium; (C) plasmid complexed with calcium phosphate incubated in D10 medium; and (D) plasmid complexed with calcium phosphate incubated in D10 medium with 25 U/mL of Pulmozyme. Each sample began with 50 ng/µL of RD114 plasmid and was analyzed on a 1.2% TAE gel.



Cost Analysis

Another consideration when assessing the feasibility of replacing Benzonase with Pulmozyme is the cost. Currently, the cost of Benzonase is $6.60 per 1000 units compared to $26.80 for 1000 units of Pulmozyme. For treating 20 liters of unconcentrated product, the cost of Benzonase (50 U/ml) is estimated at $6,600 while the cost of Pulmozyme (20 U/ml) is estimated at $10,720. To decrease costs, we currently concentrate vector product 10-fold prior to DNase treatment. Treating two liters of product decreases the cost for Pulmozyme treatment to $1,072 adding $412 to the cost of vector production when compared to the use of Benzonase. With the supply cost of a 20 liter clinical grade vector production estimated at approximately $25,000, the switch to Pulmozyme would increase the cost by approximately 1.6%.

D. Discussion

The studies presented here demonstrate that Pulmozyme can be utilized as an alternative to Benzonase for removing residual plasmid DNA from lentiviral vector products. Pulmozyme offers the advantage of being a human protein produced in Chinese Hamster Ovary (CHO) cells as opposed to the bacterially derived Benzonase. Most importantly, the availability of a pharmaceutical grade reagent will be important as vectors move from investigational agents into licensed products and the requirements for product purity escalate.

Incorporating the DNase treatment into the transfection process was evaluated as a means to decrease the time between product harvests and freeze down. While lentiviral vector half-life is longer than that of traditional gamma-retroviral vectors, efforts to insure a rapid freezing will help maximize the final vector titer (Higashikawa & Chang, 2001) and treatment at the end of transfection (Figure 5, protocol B) would minimize the time between vector harvest and freezing. The other advantage of DNase treatment in the transfection process is the subsequent media changes will greatly decrease the concentration of DNase in the final product. In contrast, treating the harvested vector leaves significant amounts of DNase in the final product (Zufferey, 2002) and exposing patients to foreign proteins has been associated with adverse events. For example, patients with pneumonia, treated with a bovine DNase, developed severe respiratory reactions due to contaminating proteins (Johnson, Goger, & Tillett, 1954; Lachmann, 1967; Raskin, 1968). Unfortunately, treatment at the

end of transfection was inferior to treating the harvested product in terms of DNA degradation.

The cause of the disparity observed between the two protocols is not readily apparent. Our findings suggest that although there does not appear to be any inhibition of the DNase by serum, there is still some degree of protection when complexed with calcium phosphate. This is apparent since Pulmozyme is able to degrade plasmid DNA complexed with calcium phosphate (Figure 9D), yet there are consistent levels of plasmid DNA detected by Q-PCR (Figures 7 and 8). This indicates that plasmid DNA complexed with calcium phosphate infers some resistance to DNase degradation, and is consistent with previously reported findings (Loyter et al., 1982). The ability to degrade the residual DNA more efficiently by treating the harvested vector is likely due to the dynamic and relatively unstable nature of the calcium phosphate complex over time (Jordan & Wurm, 2004). Whether the same degree of protection is observed with other transfection methods, such as liposome based, remains to be determined.

Although our findings indicate Pulmozyme® is a viable DNase, the timing and concentration will likely need to be tailored to the vector production method employed. Presently, there is considerable variability in large-scale lentiviral production methods, ranging from methods that provide minimal processing to those with extensive diafiltration, ion-exchange, and size separation (Merten et al., 2011; Slepushkin et al., 2003; Transfiguracion, Jaalouk, Ghani, Galipeau, & Kamen, 2003). Purification steps are likely to reduce DNA, as well as decrease the amount of residual Pulmozyme in the final product. It is possible that lower concentrations of Pulmozyme could be utilized if

subsequent purification steps contributed to plasmid DNA removal so that combined the level of DNA is decreased to undetectable levels. For clinical products, residual foreign proteins must also be minimized in order to avoid allergic reactions after vector administration. At present, the U.S. FDA has not set limits on the amount of residual DNase in a final product. The acceptable level may vary depending on whether the product is administered *ex vivo* or *in vivo*. Therefore, Pulmozyme removal should be considered in the design of downstream processing, with highly stringent processes used for vector injected directly into immune competent individuals.

The variability in production methods, along with evolving technology suggest that ongoing refinements make universal statements about DNase treatment difficult. Nevertheless, the ability to substitute Benzonase with an FDA approved reagent will improve the safety of vector products without compromising vector titer. This change will result in a minimal increase in the overall cost of clinical vector production.

III. Combining modifications to improve the utility of non-integrating lentiviral vectors

A. Introduction

Lentiviral vectors have shown success in several clinical trials, including recent gene therapy trails for the treatment of Wiskott-Aldrich syndrome (Aiuti et al., 2013) and metachromatic leukodystrophy (Biffi et al., 2013). These early clinical trials with lentiviral vectors have utilized integrating vectors which rely on integrase mediated insertion. Integrase mediated insertion by lentiviral vectors is random in nature and poses an inherent risk of insertional mutagenesis (IM) whereby a vector inserts into a loci which results in the aberrant expression of nearby genes. The risk of IM was made apparent in an early clinical trial with gammaretroviral vectors, another class of retroviruses, which resulted in adverse events leading to oncogenesis in patients after treatment (Hacein-Bey-Abina et al., 2008). This led to an increased interest in lentiviral vectors which demonstrate an improved safety profile relative to gammaretroviral vectors (Modlich et al., 2009; Montini et al., 2006). However, while clinical use of lentiviral vectors has not resulted in any oncogenic events to date, there has been an incidence of clonal expansion and dominance of hematopoietic progenitors in a human gene therapy trial for β-thalassaemia due to lentiviral vector integration triggering aberrant splicing of the *HMGA2* gene (Cavazzana-Calvo et al., 2010).

The inherent risk of IM associated with integrase mediated insertion has compelled the development of gene therapy methods which circumvent the necessity

for integrase mediated insertion. A promising avenue for safer gene therapy has been the advent of site-directed integration systems. Some of the more successful of these systems include engineered zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeats (CRISPR)-Cas systems (Joglekar et al., 2013; Lombardo et al., 2011; Lombardo et al., 2007; Mali, Esvelt, & Church, 2013; Osborn et al., 2013). However, these systems rely upon efficient delivery mechanisms such as lentiviral vectors to get into their target cells and/or provide templates for homologous recombination (HR) and require only transient expression of the components.

In order to provide a safer lentiviral vector where integrase mediated insertion is not required or warranted, non-integrating lentiviral vectors (NILV) have been developed. These vectors can provide transient expression in dividing cells and persistent episomal expression has been demonstrated for up to 9 months in non-proliferating post-mitotic cells (Yanez-Munoz et al., 2006). NILV offer many of the advantages of lentiviral vectors such as a large packaging capacity, low immunogenicity and the ability to be pseudotyped, allowing for cell-type specific targeting not possible when using other methods such as plasmids. Lentiviral vectors can be induced to be non-integrating by incorporating several mutations. These include those that inhibit the catalytic activity of the integrase protein, mutating the integrase DNA attachment sites within the U3 and U5 regions of the LTRs, and by inducing the preferential production of 1-LTR circular episomes and limiting the linear 2-LTR episomal form preferred for integrase mediated insertion by deletion of the 3’ polypurine tract (PPT) (Kantor et al.,

2011). Modifications to the integrase DNA attachment sites have consisted of single or double base-pair mutations within the 12 base-pair U3 or 11 base-pair U5 regions. However, the effectiveness of these mutations proved to be inferior as compared to point mutations to the integrase protein. In an attempt to improve the functionality of integrase DNA attachment site mutations we have evaluated a novel mutation developed by a previous student in our lab, Guiandre Joseph (GJ), by ablating the entire U3 and U5 attachment sites. The rationale behind complete ablation, as opposed to point mutations, was to inhibit any latent binding to the attachment sites and to prevent the incidence of reversion mutations. Our previous analysis of vectors with the ablated U5 attachment site revealed loss of transgene expression. After further analysis we determined that the U5 attachment site was immediately adjacent to, and overlapping a reverse transcription primer binding site (PBS). Ablation of the U5 attachment site likely resulted in inhibition of first strand synthesis during reverse transcription and loss of viable transcripts. In contrast, vectors containing the ablated U3 attachment site demonstrated significant and persistent transgene expression for up to three weeks post-transduction and this mutation was utilized for subsequent studies.

While integrase, attachment site, and PPT mutations can independently reduce the frequency of integration by several logs relative to integrating lentiviral vectors, they still demonstrate a low level of illegitimate integration. Combining mutations for inhibiting integration could help to reduce this frequency of illegitimate integration by NILV. Early studies combining integrase defective mutants with single or double base-pair mutations to the conserved terminal CA dinucleotides within the LTR

integrase DNA attachment sites (Apolonia et al., 2007; Nightingale et al., 2006) found no synergistic or additive effects. Recent studies evaluating the effectiveness of combining integrase and PPT mutant vectors have demonstrated more promising results. In two separate studies they were able to demonstrate that these elements acted independently and provided a threefold reduction in the frequency of illegitimate integration (Kantor et al., 2011; Tareen et al., 2014) as compared to the integrase mutation alone. However, given that gene therapy trials often treat in excess of 5 X 10^8 cells per patient, a 3 fold reduction in integration may not provide a clinically significant improvement to safety when using these vectors.

As lentiviral vectors are entering the clinic and new technologies are being developed which incorporate them as delivery vehicles, it is imperative to provide the safest vector possible. To improve the utility of NILV for clinical applications it will be necessary to minimize illegitimate integration and improve transgene expression. As non-integrating lentiviral vectors have demonstrated a propensity for reversion mutations (Yanez-Munoz et al., 2006), a complete ablation of the integrase attachment sites as opposed to previously evaluated point mutations could permit a reduction in the frequency of illegitimate integration when combined with other modifications. NILV have also demonstrated reduced levels of transgene expression relative to integrating vectors. Recent studies have successfully evaluated suppression of trans- and cis- acting inhibitors to episomal transgene expression. In order to improve or maintain an adequate level of transgene expression when combining modifications to reduce illegitimate integration we have evaluated incorporating a large deletion of the U3

region within the LTRs. This modification has been shown to increase transgene expression by nearly threefold by removing cis-acting inhibitors to episomal expression (Bayer et al., 2008).

Here, we present an analysis of our novel U3 LTR integrase attachment site deletion mutation and the efficacy of combining it with an integrase D116N catalytic core mutation and a 3' PPT deletion to reduce illegitimate integration and large U3 deletion of cis-acting inhibitors to improve episomal transgene expression. We report that, independently, our novel attachment site mutation does not reduce illegitimate integration as significantly as an integrase catalytic core mutation, however, combining the LTR attachment site and integrase mutations provides a significant reduction in illegitimate integration as compared to either mutation independently. Furthermore, the addition of a 3' PPT deleted mutation demonstrates an additional reduction in the frequency of illegitimate integration and the inclusion of a large U3 deletion results in significantly increased levels of transgene expression. Finally, we report that vectors with multiple modifications demonstrate transduction efficiency and stability similar to the unmodified parental vector. This provides support to their effectiveness as non-integrating vectors and the additive effects among the modifications to reduce illegitimate integration.

B. Methods

Cell culture and reagents

The production of lentiviral vector containing supernatant and the integration and expression analysis of lentiviral vector preparations was performed in HEK 293T and HEK 293 cells, respectively. Both HEK cell lines were obtained from American Type Culture Collection (Manassas, VA) and cultured in Dulbecco's modified Eagle's medium with Glutamax (Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (FBS, HyClone, Logan, UT) and 100 units/ml penicillin-streptomycin (Pen-Strep, Invitrogen).

Designing modified lentiviral vectors

Modifications to vector design were incorporated into the normal integrating lentiviral vector transfer plasmid pcDNA-CS-CGW, provided by Phillip Zoltick, and the packaging plasmid pMDL (Cell Genesys, San Francisco, CA) (Figure 10). The point mutation IN/D116N was introduced by former student GJ into pMDL by site directed mutagenesis for the production of integrase defective non-integrating lentiviral vectors (NILV). Deletion of the 12 base-pair integrase DNA attachment site within the U3 region of the transfer plasmid (LTRΔatt) was performed by GJ via sub-cloning the 3' LTR into a cloning vector by standard restriction site methods, followed by site directed mutagenesis to produce integration defective NILV. A deletion of the transfer plasmids 3' polypurine tract (ΔPPT) and a 274 base-pair portion of the LTR's U3 region from base-

pair 42-316 (ΔU3) were incorporated by standard restriction site insertion of the modified regions synthesized by GeneArt (Life Technologies, Grand Island, NY). All vector supernatants were produced by calcium phosphate transient transfection of HEK 293T cells using a 3rd generation 4-plasmid system consisting of the transfer and packaging plasmids described above, along with a Rev expression plasmid (pRSV-rev) and the VSVG envelope expression plasmid (pMDG, Cell Genesys, San Francisco, CA).

Physical titers

The physical titer of lentiviral vector supernatant was determined for normalization of vector concentration prior to transduction. Physical titer was obtained using a commercially available p24gag ELISA kit (Beckman Coulter) to measure p24 antigen using serial dilutions of vector supernatant post-transfection. The DNA titer was measured post-transduction of HEK 293 cells in order to determine vector copy number by reverse transcription Q-PCR analysis. Q-PCR was performed for detection of psi-gag sequences as previously described (Sastry et al., 2003).

Integration analysis

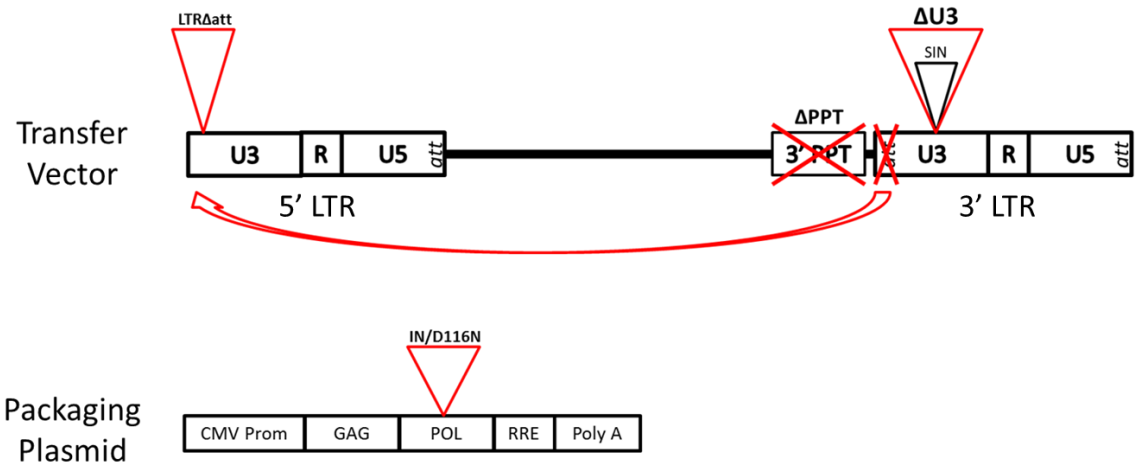
The frequency of integration was determined using an antibiotic resistant colony formation titer assay. The optimal dosage of the antibiotic Zeocin™ (Life Technologies, Carlsbad, CA) for use in our assay was determined to be 500 µg/ml using HEK 293 cells. To determine the vector infectious titer, HEK 293 cells were transduced with 8 µg/ml of polybrene (Sigma-Aldrich, St. Louis, MO) and serial dilutions of viral

supernatant containing a Bleomycin resistance transgene normalized by p24 content as described above. Selection media was added 48 hours post-transduction and changed approximately every 48 hours. The transduced cells were incubated for up to three weeks post-transduction to allow for loss and/or dilution of episomal vector to cell division, selecting for colonies with integrated vector. Viable colonies with integrated vector were then counted to calculate infectious titer based on the corresponding dilution by washing in PBS and staining with methylene blue in methanol.

Expression analysis

The level of transgene expression in transduced cells was measured by flow cytometry for GFP expression as previously described(Sastry et al., 2002) using a FACSCalibur APC or LSRII cytometer (Becton-Dickinson, San Jose, CA) and FlowJo analysis software (Tree Star, Inc., Ashland, OR). The level of transgene expression was determined by comparing the percent of GFP positive cells and median fluorescence intensity (MFI) of mutant vectors to a normal integrating control vector.

Figure 10. Schematic demonstrating several modifications incorporated into NILV design. Mutations and deletions incorporated in novel NILV intended to reduce illegitimate integration and improve transgene expression. Modifications are illustrated in red. The transfer vector presents three modifications: a deletion of the 3' polypurine tract (ΔPPT) at the terminal end of the vector backbone, immediately adjacent to the 3' LTR; a deletion of the U3 integrase attachment site (*att*) in the 3' LTR which is copied to the 5' LTR following reverse transcription as indicated by the red arrow (LTRΔatt); and a large deletion of 274 base-pairs in the U3 region (ΔU3) in addition to the previous self-inactivation deletion (SIN) included in 3rd generation LV. The packaging plasmid demonstrates a mutation to the pol gene region (IN/D116N) in the catalytic core domain of integrase.



C. Results

Novel U3 LTR attachment site deleted NILV have a reduced frequency of integration

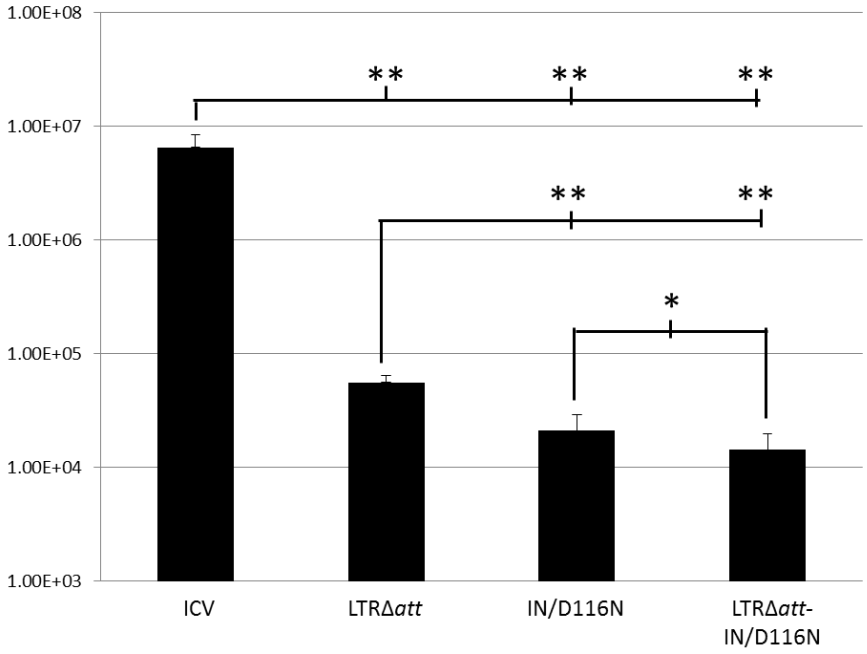
The utility of a NILV is almost entirely dependent upon its efficiency of inhibiting or minimizing the frequency of integration. It is therefore pertinent that when developing NILV for clinical use, that any illegitimate integration is thoroughly evaluated. This is necessary in order to provide for informed risk assessments when using these vectors for transient or sustained episomal expression and where insertional mutagenesis is a concern.

The frequency of integration by a novel U3 LTR integrase attachment site deleted NILV (LTRΔatt) expressing a bleomycin resistance transgene was assessed in three separate experiments by an antibiotic colony formation assay following drug selection of transduced HEK 293 cells. In Figure 11 the efficiency of the novel NILV, LTRΔatt, to reduce integration frequency was compared relative to an integration competent vector (ICV), an NILV with a point mutation in the catalytic core domain of integrase (IN/D116N) and an NILV containing both modifications to the attachment site and integrase (LTRΔatt-IN/D116N).

The three NILV all reduced the frequency of integration approximately two logs below that of an ICV. Similar to LTR integrase attachment site mutant vectors with single or double base-pair mutations reported in the literature, our novel vector (LTRΔatt) integrated at a significantly higher frequency than the more commonly used integrase catalytic core mutant (IN/D116N). Interestingly, in contrast to other reports

when combining integrase and attachment site mutations (Apolonia et al., 2007; Nightingale et al., 2006), the double mutant (LTRΔatt-IN/D116N) significantly reduced the frequency of integration by up to 4 fold as compared to either mutation independently. The complete ablation of the U3 LTR integrase attachment site demonstrates a synergistic reduction in the frequency of illegitimate integration when combined with the integrase catalytic core mutation. These findings laid the groundwork for investigating the combining of multiple modifications to improve the efficacy of NILV.

Figure 11. A deletion of the U3 integrase attachment site can effectively reduce the frequency of integration in an NILV. Average antibiotic resistant colony formation titer assay results from two independent experiments comparing the frequency of integration among an integration competent vector (ICV), a U3 LTR integrase attachment site mutant NILV (LTRΔatt), an integrase deficient NILV (IN/D116N), and a double mutant NILV (LTRΔatt-IN/D116N). Y-axis represents infectious units/ml of vector supernatant. Error bars indicate standard deviation of the mean. Asterisk indicate a significant difference between vectors. * - $p < 0.05$; ** - $p < 0.005$



Multiple modifications to vector design can be combined to improve the utility of NILV

It has been reported that point mutations to the LTR integrase attachment sites have not provided as significant of a reduction in the frequency of integration as compared to integrase catalytic core mutations, similar to the results with LTRΔatt. It has also been reported that combining LTR attachment site mutations with integrase mutations provided no further reduction in the frequency of integration as compared to using the integrase mutants alone (Apolonia et al., 2007; Nightingale et al., 2006). The finding of reduced integrations when LTRΔatt was combined with the D116N integrase mutation encouraged us to evaluate other modifications that could further improve NILV.

Two additional modifications to the transfer vector plasmid reported in the literature were investigated in hopes of further improving NILV function. These include a deletion of the 3' PPT to decrease integration by promoting the formation of 1-LTR circles and reducing linear episomes (Kantor et al., 2011). While the above modifications were aimed at decreasing integration, we also evaluated a 273 base-pair deletion in the U3 region of the LTR that removes cis-acting inhibitors and has the potential to increase transgene expression. The mechanism by which these sequences inhibit transgene expression is unclear, however, early studies have indicated their existence (Bayer et al., 2008; Hoover et al., 1996; Y. Lu, Stenzel, Sodroski, & Haseltine, 1989; Rosen, Sodroski, & Haseltine, 1985). The ΔU3 and ΔPPT modifications were incorporated into either an integration competent vector (CS-CZW) or the attachment site mutant (LTRΔatt) transfer vector plasmid and combined with either an integrase-competent (pMDL) or

integrase defective (IN/D116N) packaging plasmid. The resulting plasmids were verified by sequencing analysis and p24 ELISA analysis of vector supernatant to ensure the modifications did not adversely affect vector titer (data not shown). Each vector modification was evaluated independently and in combination with each variation for a total of 16 vectors containing either a bleomycin resistance or GFP transgene for assessing integration or expression, respectively (Table 1).

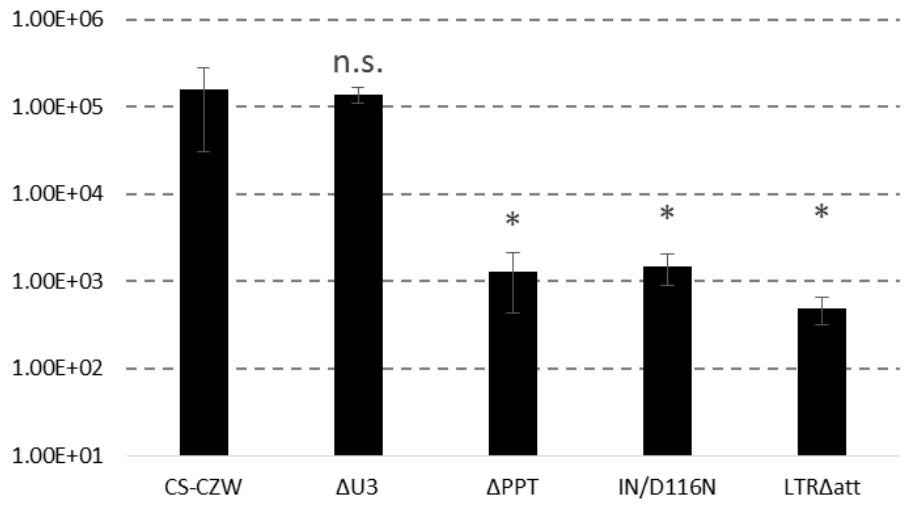
Combining multiple modifications for inhibiting vector integration can reduce the frequency of illegitimate integration by NILV

In order to determine the efficacy of modified NILV, we first sought to evaluate the frequency of integration utilizing vectors containing a bleomycin resistance transgene in a quantitative colony formation assay. We first evaluated the individual mutations, that each inhibit a different point in the vector life cycle, in order to determine their efficacy independently and to provide a baseline on their integration frequency (Figure 12). The ΔU3 modification had no effect on integration and performed similar to the ICV (CS-CZW). The NILV modifications (ΔPPT, IN/D116N and LTRΔatt) exhibited up to a three log reduction in integration frequency. Individually, vectors with a single modification integrated at an expected frequency relative to results in the literature (Kantor et al., 2011; Negri et al., 2007)

Table 1. Combining four modifications to lentiviral vector design results in 16 viable vectors for integration and expression analysis. Representation of the 16 different lentiviral vector combinations possible when combining transfer vector modifications with a packaging plasmid mutation. Shaded areas indicate the modifications included per each vector combination in each row. CS-CZW – Normal integration competent transfer vector backbone containing either a bleomycin as shown or GFP transgene (CS-CGW); LTRΔatt – LTR U3 integrase attachment site deletion; ΔPPT – 3' polypurine tract deletion; ΔU3 – 273 base-pair U3 deletion; pMDL – packaging plasmid with an intact integrase; IN/D116N – a integrase catalytic core point mutation at residue D116.

Transfer Vector Plasmid				Packaging Plasmid	
CS-CZW	LTRΔatt	ΔPPT	ΔU3	pMDL	IN/D116N
■				■	
■					■
■		■		■	
■		■			■
■			■	■	
■			■		■
■		■	■	■	
■		■	■		■
	■			■	
	■				■
	■	■		■	
	■	■			■
	■		■	■	
	■		■		■
	■	■	■	■	
	■	■	■		■

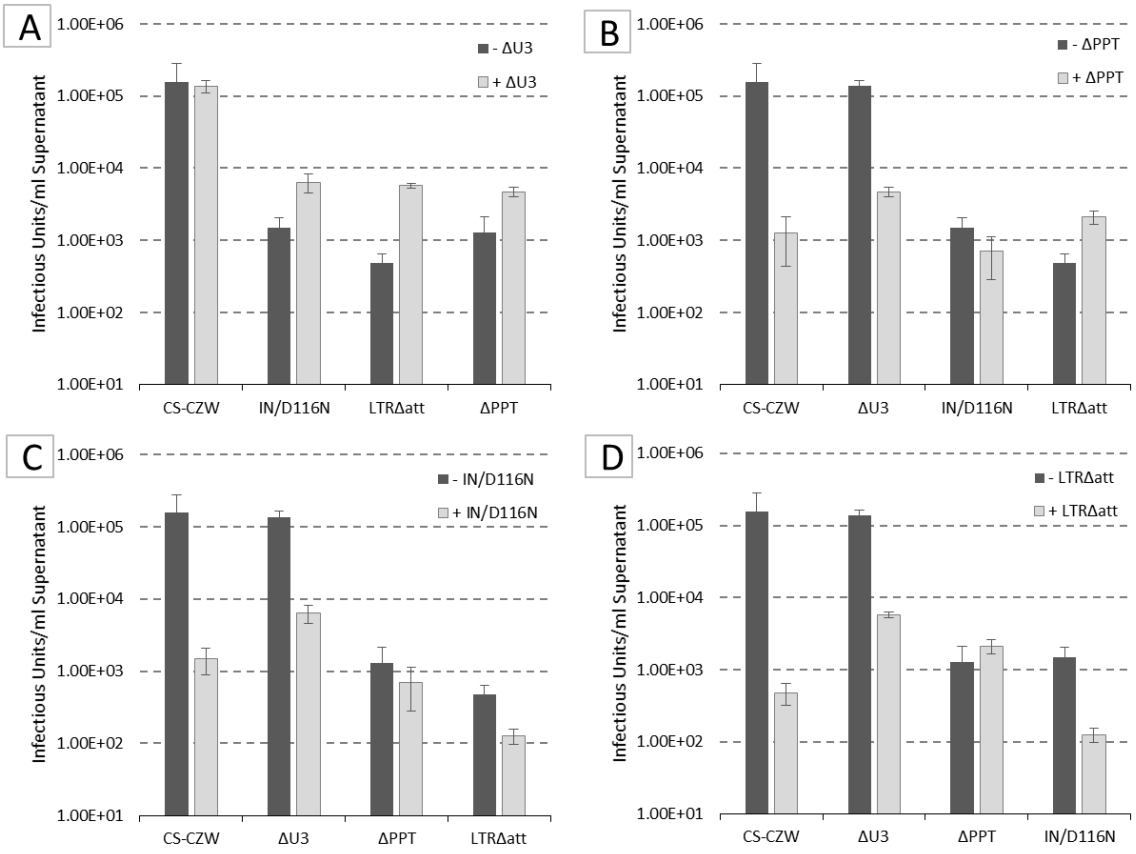
Figure 12. The frequency of integration can be reduced independently in NILV by incorporation of mutations that inhibit integration at different points in the viral life cycle. Average bleomycin resistant colony formation results from two experiments comparing the frequency of integration among individual mutations to lentiviral vector design. From left to right the vectors analyzed are: an integration competent vector (CS-CZW), an integrating vector with a large deletion of cis-acting inhibitors to episomal transgene expression in the U3 region of the vectors LTRs (ΔU3), an NILV with a deletion of the vectors 3' polypurine tract (ΔPPT), an integrase deficient NILV (IN/D116N), and a U3 LTR integrase attachment site mutant NILV (LTRΔatt). Y-axis represents infectious units/ml of vector supernatant. Error bars indicate standard deviation of the mean. Asterisk indicate a significant difference relative to the integration competent vector (CS-CZW). * - $p < 0.05$; n.s. – not significant.



When pairing the modifications with each other the expected results were also observed. Combining each NILV vector with the ΔU3 modification (Figure 13A) resulted in an increase in the frequency of integration, indicating that removing cis-acting inhibitors to episomal transgene expression can result in more viable episomes which can then provide substrates for illegitimate integration.

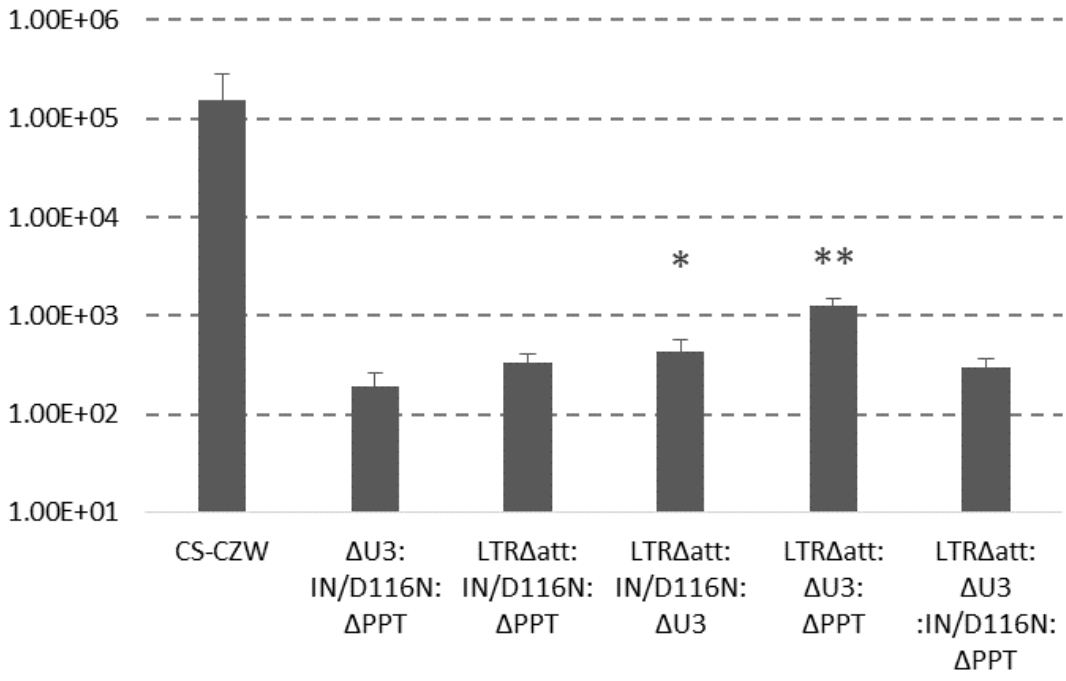
Pairing each NILV mutation with the ΔPPT modification provided modest reductions in integration frequency (Figure 13B). Similar to results reported by Kantor et al. and Tareen et al., pairing ΔPPT with IN/D116N provided a further reduction in integration as compared to either mutation individually, indicating that each modification works independently to inhibit integration (Kantor et al., 2011; Tareen et al., 2014). Interestingly, combining the ΔPPT modification with LTRΔatt resulted in a frequency of integration similar to that of the ΔPPT alone, with no further reduction in integration frequency. This may be due to the fact that the two modifications are immediately juxtaposed to one another in the vector possibly providing redundant mechanisms of inhibiting integration. In previous work by GJ, analysis of episomal forms indicated a persistent predominance of 1-LTR circles relative to linear and 2-LTR forms by LTRΔatt, similar to effect observed when deleting the 3' PPT to induce the preferential formation of 1-LTR circles. In contrast, combining the IN/D116N mutation with LTRΔatt provided the greatest reduction in integration frequency (Figure 13C, D). This is likely due to an additive effect between two independent non-redundant mechanisms for inhibiting integration.

Figure 13. The frequency of integration can be reduced in NILV by pairing modifications to vector design for inhibiting vector integration. Average bleomycin resistant colony formation assay results from two experiments comparing the frequency of integration among paired mutations to LV design. The horizontal axis designates the vector or independent modification (dark gray bars) and pairing it with the indicated mutation (light gray bars). CS-CZW - an integration competent vector; (A) ΔU3 modification - a vector with a U3 deletion of cis-acting inhibitors to transgene expression; (B) ΔPPT modification- an NILV with a deletion of the vectors 3' polypurine tract; (C) IN/D116N modification - an integrase deficient NILV; (D) LTRΔatt modification - a U3 LTR integrase attachment site mutant NILV. Error bars indicate SD of the mean.



Next, we evaluated the efficacy of combining three or more of the modifications to vector design in order to elucidate their combined effect on integration frequency. Combining three or more of the modifications to vector design provided the most consistent reduction in illegitimate integration frequency (Figure 14). Vector combinations with at least two non-redundant modifications for reducing integration frequency were able to efficiently inhibit illegitimate integration. The presence of the ΔU3 modification resulted in a significant increase in illegitimate integration when combined with LTRΔatt-IN/D116N and LTRΔatt-ΔPPT as compared to the other combinations in this group. This was especially evident in the vector combining the ΔU3 deletion with the possibly redundant mutations LTRΔatt and ΔPPT. However, a vector with all 3 integration inhibiting modifications, LTRΔatt-ΔPPT-IN/D116N, did not demonstrate any significant increase in illegitimate integrate when combined with the ΔU3 deletion. This demonstrates the ability of combined integrase inhibiting modifications to reduce illegitimate integration in the presence of more viable episomes.

Figure 14. The frequency of integration can be reduced in NILV by combining multiple modifications to vector design. Average bleomycin resistant colony formation assay results from two experiments comparing the frequency of integration among combining multiple modifications to lentiviral vector design. CS-CZW - an integration competent vector; ΔU3 - a U3 deletion of cis-acting inhibitors to episomal transgene expression in the LTRs; ΔPPT - a deletion of the vectors 3' polypurine tract; IN/D116N - an integrase deficient NILV modification; LTRΔatt - a U3 LTR integrase attachment site mutant NILV modification. Error bars indicate standard deviation of the mean. Y-axis indicates infectious units/ml of vector supernatant. Asterisk indicate a significant difference relative to the other NILV combinations. * - $p < 0.05$; ** – $p < 0.005$.



Removing cis-acting inhibitors to episomal transgene expression can improve the level of transgene expression when combined with multiple modifications that inhibit vector integration

An important factor plaguing the development of NILV has been significantly reduced levels of transgene expression relative to integration competent vectors (ICV). In order to analyze our modifications to vector design we utilized a GFP transgene and assessed vector expression by flow cytometry (Figure 15).

The three modifications for reducing integration frequency (ΔPPT, IN/D116N and LTRΔatt) individually produced levels of transgene expression between one and two logs below that of a wild-type vector as previously demonstrated for NILV (Bayer et al., 2008; Kantor et al., 2011; Naldini et al., 1996; Philippe et al., 2006; Vargas et al., 2004).

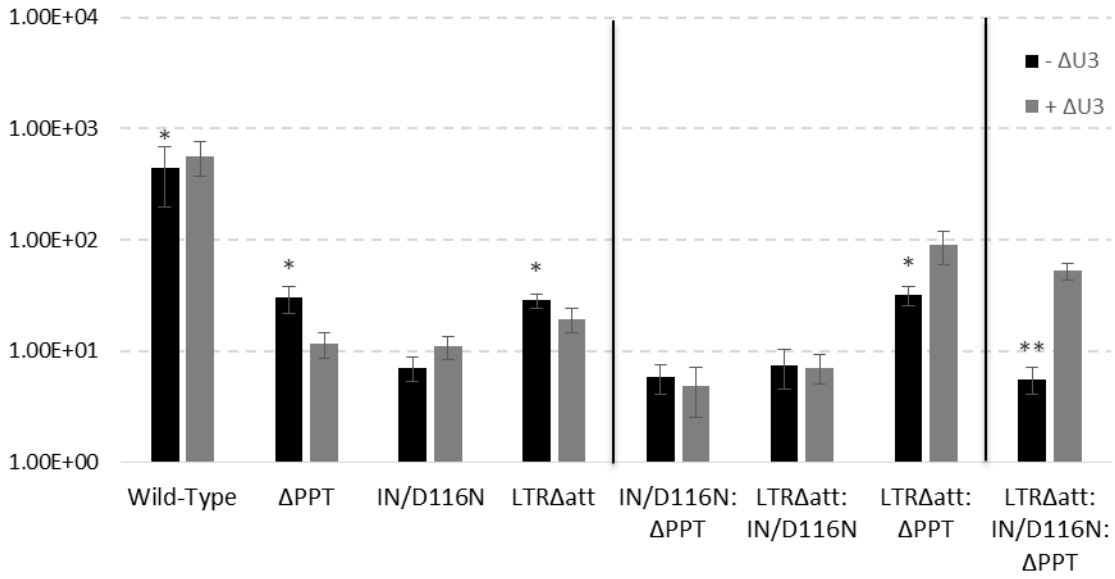
The addition of the ΔU3 modification was able to significantly improve the level of expression of a wild-type vector at 1.3X fold increased MFI. When paired with the ΔU3 deletion there was no significant improvement in expression levels with the ΔPPT or LTRΔatt modifications. However, when combining the ΔU3 with the IN/D116N modification, there was a 1.6X increase in expression likely due to more viable episomal forms in the absence of cis-acting inhibitors removed by the ΔU3 deletion.

There was no significant effect on expression when pairing the ΔPPT deletion with the integrase mutation IN/D116N, in line with results by Kantor et al. using an integrase mutant D64E NILV (Kantor et al., 2011). Pairing the IN/D116N mutation with LTRΔatt modification resulted in a significant decrease in expression. This combination demonstrated one of the lowest frequencies of integration, thus having fewer

integrated vector and less viable episomal forms in the absence of ΔU3 available to express the transgene.

Combining three or more modifications to vector design had variable results. The ΔU3 modification was unable to provide any significant improvement in expression to the IN/D116N-ΔPPT or LTRΔatt-IN/D116N pairs. However, combining the ΔU3 deletion with the LTRΔatt-ΔPPT pair was able to significantly improve the expression of this combination by 2.8X giving the best expression for any of the vector combinations. This is probably due to having more viable episomal forms and a higher frequency of illegitimate integration providing better transgene expression. Combining ΔPPT, LTRΔatt, and IN/D116N provided a reduced level of transgene expression, likely due to the combination of all three modifications for reducing integration frequency without the ΔU3 deletion to remove cis-acting inhibitors. Adding the ΔU3 deletion to this combination, creating a vector with all four modifications, was able to significantly rescue the level of transgene expression. This vector with all four modifications to vector design gave the second best level of transgene expression, but with a significantly lower frequency of integration relative to the LTRΔatt-ΔPPT-ΔU3 vector which also demonstrated significantly improved transgene expression (Figure 14).

Figure 15. Combining a ΔU3 modification with multiple modifications to reduce integration frequency can maintain and/or improve transgene expression levels. Average flow cytometry for GFP expression results from two experiments comparing the levels of transgene expression among combining modifications to lentiviral vector design with a ΔU3 deletion for removing inhibitors to transgene expression. Black bars indicate the geometric mean fluorescence intensity (MFI) for the indicated vector on the x-axis, gray bars indicate the MFI when combining the indicated vector with the ΔU3 deletion. Wild-type - normal integrating vector (CS-CGW); ΔPPT - a deletion of the vectors 3' polypurine tract; IN/D116N - an integrase deficient NILV modification; LTRΔatt - a U3 LTR integrase attachment site mutant NILV modification. Error bars indicate standard deviation of the mean. Asterisk indicate a significant difference relative to combination with the ΔU3 deletion. * - $p < 0.05$; ** – $p < 0.005$.

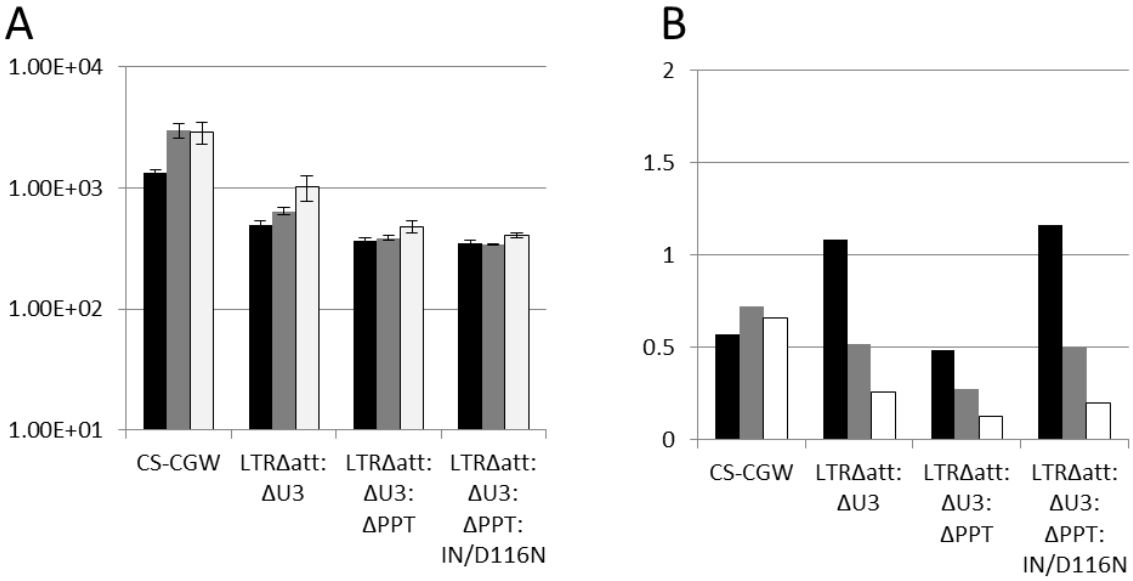


NILV with multiple modifications demonstrate expected copy number trends and demonstrate levels of transgene expression relative to their frequency of integration

In order to further verify the non-integrating nature of the NILV, copy number was evaluated by RT-Q-PCR up to 96 hours post-transduction. Since HEK 293 cells double approximately every 24 hours, an optimal NILV population would be expected to demonstrate a reduction in copy number of roughly 50% with each cell division. For this analysis three NILV with differing frequencies of integration were selected (LTRΔatt-ΔU3, LTRΔatt-ΔU3-ΔPPT and LTRΔatt-ΔU3-ΔPPT-IN/D116N). The vector combinations were compared over time relative to an integration competent vector (CS-CGW). Regardless of their frequency of integration, each NILV demonstrated the predicted reduction in copy number over time (Figure 16B).

The low level of illegitimate integration observed with these NILV would likely be unable to affect any observable increase in copy number until a much later time allowing for sufficient replication. However, the levels of transgene expression obtained by the NILV demonstrated an increase relative to integration frequency (Figure 16A). The LTRΔatt- ΔU3 NILV showed a consistent increase in expression over time, demonstrating a correlation with the higher level of illegitimate integration by this vector relative to the other NILV. This is in line with work by Pelascini et al. where episomal NILV were considered preferential targets for epigenetic silencing by chromatin-remodeling histone deacetylation (Pelascini et al., 2013), thus increased integration would provide for increased levels of transgene expression with similar copy number.

Figure 16. NILV demonstrate expected trends for copy number over time while providing levels of transgene expression relative to their frequency of integration. (A) Flow cytometry results for MFI of GFP expression at 48 (black bars), 72 (gray bars) and 96 hours (white bars) post-transduction. Error bars indicate standard deviation of the mean among three replicates per time point. (B) Vector copy number determined from RT-Q-PCR analysis measured at 48 (black bars), 72 (gray bars) and 96 hour (white bars) time points post-transduction. Horizontal axes indicate the vector mutations per vector. CS-CGW - an integration competent vector; ΔU3 - a U3 deletion of cis-acting inhibitors to episomal transgene expression in the LTRs; ΔPPT - a deletion of the vectors 3' polypurine tract; IN/D116N - an integrase deficient NILV modification; LTRΔatt - a U3 LTR integrase attachment site mutant NILV modification.



NILV demonstrate levels of transgene expression near the approximation of their theoretical maximum and relative to their frequency of integration

An inherent difficulty in analyzing NILV transduction efficiency and stability in dividing cells is the dilution of transduced cells with each cell division. This can lend to an under-estimation of NILV activity. Integrating lentiviral vectors can maintain a constant percentage of transduced cells over time by concomitant replication with each cell division. The transduction efficiency of NILV, on the other hand, is limited to the number of cells available for infection at the time of transduction due to their episomal nature and lack of origins of replication.

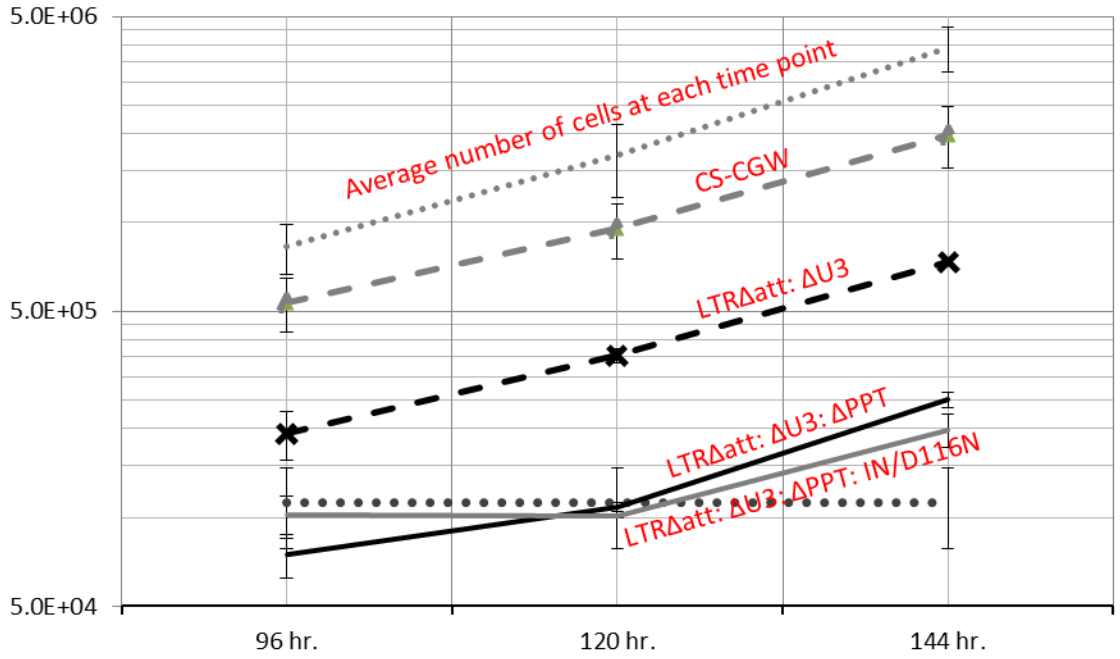
In order to determine the transduction efficiency and stability of NILV in dividing cells, one must take into account the number of cells present at transduction and their rate of replication. The initial percentage of cells containing NILV is expected to decrease by 50% with each cell division assuming no integration and one episomal copy per cell. Any significant deviation from this trend by NILV would be indicative of illegitimate integration resulting in higher titers, or loss of episomes by degradation or silencing resulting in decreasing titer.

In order to account for these characteristics of NILV, the total number of cells were counted at the day of transduction (Day 1) and 96, 120, and 144 hours post-transduction. The percentage of GFP positive cells were measured by flow cytometry at 96, 120, and 144 hours' post-transduction allowing adequate time for transgene expression. For this analysis three NILV with differing frequencies of illegitimate integration were compared relative to an ICV (CS-CGW), including LTRΔatt-ΔU3,

LTRΔatt-ΔU3-ΔPPT and LTRΔatt-ΔU3-ΔPPT-IN/D116N. For the analysis two reference lines were plotted. A line indicating the total number of cells at each time point (Figure 17, light gray, small, dotted line), which would be indicative of 100% integrated transduction efficiency. Also, a line indicating the number of cells present at transduction (Figure 17, dark gray, large, dotted line), indicative of the maximum transduction efficiency possible for an NILV without any significant frequency of illegitimate integration. The ICV demonstrated the expected expression pattern (Figure 17, light gray, dashed line), maintaining an approximately constant percentage of transduced cells over time relative to their initial efficiency of infection and the number of cells at each time point. The vector LTRΔatt-ΔU3, the NILV analyzed with the highest frequency of illegitimate integration, demonstrated a similar trend albeit with a significantly reduced percentage of transduced cells (Figure 17, black, dashed line). LTRΔatt-ΔU3-ΔPPT gave an initial transduction efficiency at 72 hours post-transduction of 67% relative to the number of cells available at the time of transduction (Figure 17, solid, black line). The percentage of transduced cells remained below the theoretical maximum at 96 hours while slowly increasing until after 120 hours where they exceeded this maximum. This is possible due to its lower frequency of illegitimate integration requiring additional time to replicate and begin to increase the titer to an observable level. However, the percentage of transduced cells relative to the total number of cells at this time point is very low at 6%, leaving room for experimental error due to the limitations of the FACS machine and possible false positives. The vector with the lowest frequency of integration analyzed, LTRΔatt-ΔU3-ΔPPT-IN/D116N, showed a similar trend

while remaining at more consistent level below the theoretical maximum and did not exceed it until a later time point (Figure 17, solid, gray line). This demonstrates the significantly lower frequency of integration by this vector, indicating stable episomal expression over time with minimal effects observable due to illegitimate integration. These four vectors demonstrate the improved efficiency of combining multiple modifications to reduce illegitimate integration providing a safer and more effective tool for gene therapy applications.

Figure 17. NILV demonstrate transduction efficiencies representative of their frequency of illegitimate integration. Cell counts (dotted lines) and flow cytometry results for infectious titer based on the percentage of GFP positive cells. Large, dark gray dotted line denotes the average number cells at the day of transduction as a reference for NILV transduction efficiency. Small, light gray dotted line represents the average number of cells at each time point analyzed as a reference for integration competent vector transduction efficiency and illegitimate integration by NILV. Solid and dashed lines indicate average titer results. Gray dashed line –integration competent vector (CS-CGW); Black dashed line – an NILV with a high frequency of illegitimate integration (LTRΔatt-ΔU3); Solid black line – NILV (LTRΔatt-ΔU3-ΔPPT) with the highest MFI; Solid gray line – NILV (LTRΔatt-ΔU3-ΔPPT-IN/D116N) with the optimal ratio of transgene expression and reduced illegitimate integration. Error bars indicate standard deviation of the mean for three replicates at each time point.



D. Discussion

The goal of this work was to evaluate a novel NILV and assess the value of combining it with multiple modifications to vector design in order to improve their safety and utility for gene therapy applications. The evaluation of a novel NILV (LTRΔatt), with a deletion of the 12 base-pair integrase DNA attachment site within the U3 region of the 3' LTR, demonstrated a significant reduction in the frequency of integration relative to an ICV. However, while LTRΔatt provided a significant reduction in the frequency of integration it proved inferior to reduction in illegitimate integration achieved by an integrase mutant NILV, similar to what others have observed when evaluating point mutations within the integrase DNA attachment sites (Apolonia et al., 2007; Masuda et al., 1995; Nightingale et al., 2006). Of greater interest, this initial work demonstrated an improvement in the reduction of integration frequency when combining this novel NILV modification with a mutation in the catalytic core domain of integrase (IN/D116N). This finding suggested that combining multiple modifications to inhibit integration can improve the safety profile of NILV by significantly reducing the frequency of illegitimate integrations.

To further investigate these findings, a deletion of the vector's 3' polypurine tract (ΔPPT), intended to further reduce the frequency of integration, was explored along with a large deletion within the U3 region of the vector's LTRs (ΔU3) aimed at improving the level of transgene expression in NILV. These additional modifications were combined with the LTRΔatt and the IN/D116N modifications to produce two sets

of 16 different vectors with either a GFP or Bleomycin resistance transgene. Our integration analysis results demonstrated a synergistic effect upon combining modifications to reduce integration providing the groundwork for combining other modifications which may further improve the profile of NILV. As NILV have historically demonstrated a reduced level of transgene expression relative to ICV, the combination of multiple modifications to reduce integration posed the risk of a further reduction to expression levels. In order to compensate for the possibility of this phenomenon, we evaluated the addition of a ΔU3 deletion of cis-acting inhibitors to transgene expression. We demonstrated that incorporation of the ΔU3 deletion was able to rescue the level of transgene expression in NILV with combinations of multiple modifications to reduce illegitimate integration. These results demonstrate that combining NILV modifications can improve their transgene expression and safety profile.

We then sought to further confirm the integration deficient nature of the NILV combinations with a more in depth analysis. We chose three NILV for analysis that displayed optimal levels of transgene expression and varying frequencies of illegitimate integration in order to elucidate how these factors affect their behavior. We found that, regardless of integration frequency, the NILV all demonstrated reductions in copy number over time as expected in actively dividing cells. However, as NILV transgene expression levels are expected to remain constant or possibly decline over time, vectors with increased frequencies of illegitimate integration would demonstrated relative trends of increased transgene expression due to integrated vector replicating with each cell division.

Finally, we took an in depth look at the transduction efficiency of NILV by comparing the percentage of transduced cells with the number of cells present over time. We were able to demonstrate that NILV are able to achieve comparable levels of transduction efficiency relative to ICV. Our results indicate that our novel NILV combinations are able to efficiently transduce cells and provide levels of transgene expression and rates of integration superior to current NILV.

IV. Modified insertion site analysis facilitates the elucidation of a unique mode of insertion presented by a novel U3 LTR integrase attachment site deleted non-integrating lentiviral vector mutation

A. Introduction

The use of viral vectors for clinical applications has been steadily increasing as they have been continually optimized and demonstrate a safer profile relative to early viral vectors. As new systems are developed that may provide potentially increased safety, the components of these systems will require intense scrutiny in order to assess their applicability in patients. Site-directed integration systems, including engineered zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeats (CRISPR)-Cas systems (Joglekar et al., 2013; Lombardo et al., 2011; Lombardo et al., 2007; Mali et al., 2013; Osborn et al., 2013), have recently demonstrated promising clinical utility. These systems can improve the safety of gene therapy by omitting the necessity for viral-mediated integration and the risk of insertional mutagenesis and facilitate insertion into or correction of specific loci. However, systems such as these still require efficient mechanisms for introduction into a target cell and trafficking to the nucleus. Lentiviral vectors are an effective gene therapy tool for this application with their large packaging capacity, broad tropism through pseudotyping, and their ability to transduce non-dividing cells, among other beneficial characteristics. However, wild-type lentiviral vectors with integrase-mediated insertion can detract from the advantages of using site-

directed integration. While lentiviral vectors have not demonstrated any severe adverse effects as observed with early gamma retroviral vectors due to insertional mutagenesis (Braun et al., 2014; Hacein-Bey-Abina et al., 2008; Howe et al., 2008; Ott et al., 2006), the risk will persist with the random nature of integrase-mediated insertion. This was made evident in a recent gene therapy trial for β-thalassaemia where integrase-mediated lentiviral vector insertion resulted in a dominant cell clone due to transcriptional activation of HMGA2 (Cavazzana-Calvo et al., 2010).

Non-integrating lentiviral vectors (NILV) provide a viable alternative to integration-competent viral vector systems, with a reduced risk of insertional mutagenesis while retaining many of the benefits of lentiviral vectors. NILV can provide for transient expression of site-directed integration systems and/or their templates for homologous recombination, reducing the risks associated with persistent expression and the odds of off-target insertions. Other clinical applications for NILV include vaccination and cytotoxic cancer therapies, among others. However, while NILV can provide a safer alternative to integration competent vectors, they still retain a low level of illegitimate integration which must be accounted for. We and others have evaluated combining modifications to reduce the frequency of illegitimate integration in NILV, however it is unlikely that a complete ablation of illegitimate integration is feasible. The majority of illegitimate integrations by NILV appear to occur due to non-integrase-mediated methods of insertion (Gaur & Leavitt, 1998; Matrai et al., 2011; Nightingale et al., 2006), similar to integrations observed with other non-integrating systems at sites of chromosomal breakage through homologous recombination or non-homologous end-

joining pathways (Koyama et al., 2013; D. G. Miller et al., 2004). In order for NILV to meet stringent requirements for progression into the clinic, it will be necessary to thoroughly scrutinize any illegitimate integrations. This will require methods to accurately analyze these insertion sites during development and for follow up of patients post gene therapy treatment as set for in FDA guidelines (Guidance for Industry: Gene Therapy Clinical Trials – Observing Subjects for Delayed Adverse Events).

Techniques previously available for the analysis of vector insertion sites include ligation-mediated PCR (LM-PCR) (Smith, 1992), linear amplification-mediated PCR (LAM-PCR) (Schmidt et al., 2007), and non-restrictive LAM-PCR (nrLAM-PCR) (Gabriel et al., 2009). While these methods can accurately identify the majority of vector insertion sites when analyzing integration competent vectors, these methods lack the sensitivity to account for the variability observed in NILV insertion sites. These methods rely upon intact vector LTRs for the capture and amplification of vector insertion sites. The majority of NILV, which integrate via non-integrase mediated insertion, have presented large deletions and/or insertions at the vector-genome junction including large truncations or ablations of the vectors LTR (Koyama et al., 2013; Matrai et al., 2011; Nightingale et al., 2006) which could result in the loss of capture by these methods. Techniques for accurately analyzing NILV insertion sites have previously employed the use of cloning vectors in order to assess the features of the vector-genome junction. These techniques can be very labor and time intensive and are not feasible for any high-throughput analysis. Recently developed techniques for analysis of vector insertion sites have utilized sonication to shear the genome of transduced cells,

followed by capture and amplification of fragments containing vector-genome junctions (De Ravin et al., 2014; S. Zhou et al., 2014). However, these techniques also rely upon intact LTR sequences in order to capture the fragments containing vector insertion sites. In order to evaluate NILV insertion sites with sufficient sensitivity and in a high-throughput manner, existing techniques will need to be modified or new techniques developed in order to facilitate their analysis.

In this study, we sought to evaluate the insertion sites of clonal populations transduced by either a wild-type lentiviral vector, a novel LTR integrase attachment site mutant NILV (LTRΔatt), an integrase deficient NILV (IN/D116N), or a double mutant containing both mutations (LTRΔatt-IN/D116N). We have previously demonstrated that our novel LTRΔatt vector provides a significantly higher frequency of illegitimate integration as compared to IN/D116N and that combining the mutations appears to result in a further reduction in integration frequency as compared to either mutation independently. We sought to analyze the vector-genome junctions of these vectors in order to elucidate the mechanisms behind their illegitimate integration which may lend to explaining the disparity observed in the frequency of integration with our novel LTRΔatt vector. To capture and analyze the vector insertion sites we employed LAM-PCR utilizing either 454 pyrosequencing or Illumina MiSeq platform. We also adapted protocols for insertion site capture utilizing sonication-mediated genomic fragmentation and Illumina MiSeq sequencing in an attempt to improve the capture and analysis of NILV insertion sites. We demonstrate that a reliance upon intact LTR fragments in methodologies for vector insertion site capture and amplification can hamper the ability

to evaluate NILV insertion sites with their variable vector-genome junction features. Accounting for the possibility of LTR truncations in protocols for NILV insertion site capture provides for improved analysis.

B. Methods

Isolating Clonal DNA

Individual clones were selected by transducing HEK 293 cells with serial dilutions of vector supernatant containing a Bleomycin resistance transgene. The vectors for transduction included an integration competent vector (CS-CZW), an integrase catalytic core mutant NILV (IN/D116N), a novel LTR U3 integrase attachment site deleted mutant NILV (LTRΔatt), and a double mutant containing both the integrase and attachment site mutations (LTRΔatt –IN/D116N). Transduced cells were expanded in Zeocin™ (Life Technologies, Carlsbad, CA) containing selection media for up to three weeks post-transduction. Individual colonies were then isolated with cloning disks and expanded. Genomic DNA from the expanded clones was extracted using the Puregene DNA Purification Kit (Gentra Systems, Inc., Minneapolis, MN).

Insertion Site Capture

For the first two experiments the vector-genome junctions were amplified by linear amplification mediated PCR (LAM-PCR) off of the 3' LTR as previously described (Schmidt et al., 2007). Briefly (Figure 18), 200 ng of genomic DNA was amplified with two rounds of 50 cycle linear PCR using a biotinylated LTR-specific primer. The single stranded PCR products were captured with streptavidin coated magnetic beads and double stranded using a Klenow enzyme and hexanucleotide primers. The bound, double stranded products were digested with the restriction enzyme *Tsp*509I (New

England Biolabs) and a linker cassette was ligated onto the digested sticky end. Nested exponential PCR was then performed using LTR-specific and linker cassette-specific primers for the first round of PCR. During the second PCR iteration of the nested PCR, primers containing sequencing adapters and multiplex identifiers/indexes were incorporated for Roche 454 or MiSeq library preparation.

For our third experiment the vector-genome junctions were amplified by modified techniques adapted from Ravin, et al and Zhou, et al (De Ravin et al., 2014; S. Zhou et al., 2014). Briefly, to perform the modified technique for surveying vector insertion sites the genomic DNA was first sheared to an approximate size of 1,000 bp using Covaris sonication (Covaris, Wohurn, MA). Fragment size was evaluated using the Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, CA). Starting with 1000 ng of sheared fragments, the ends were repaired using the NEBNext End Repair Module (NEB, Ipswich, MA) and 3’ dA-tailed with the NEBNext dA-Tailing Module (NEB, Ipswich, MA). Products greater than ~300 bp were selected and purified using a 0.8:1 ratio of AMPure XP beads (Beckman Coulter, Brea, CA) to dA-tailed product. A partially double stranded linker cassette with a 5’ T-overhang was then ligated to the ends of the purified products using the Fast-Link DNA Ligation kit (Epicentre, Madison, WI; LamTlinkerL: 5’ GACCCGGGAGATCTGAATTCAGTGGCACAGCAGTTAGGT 3’; LamTlinkerS: 5' CCTAACTGCTGTGCCACT 3’). Exponential PCR was then performed using the ligation product as a template with a biotinylated vector-specific primer for the 5’ or 3’ end of the vector and a linker-specific primer (LentiLAM-LC1: 5’ GACCCGGGAGATCTGAATTC 3’). Vector-specific PCR products were then enriched using Streptavidin-coated M-280

Dynabeads (Invitrogen, Carlsbad, CA). Next, nested PCR using a second MiSeq indexed LTR-specific primer for the 5’ or 3’ end and another linker-specific primer (LentiLAM-LC2: 5’ AGTGGCACAGCAGTTAGG 3’) was performed using the enriched product as a template. The final PCR products were then purified using a 1.8:1 ratio of AMPure XP beads and quantified using a Qubit Fluorometer (Life Technologies, Grand Island, NY).

5’ LTR Vector-Genome Amplification for 3’ LAM-PCR Verification

In order to confirm the vector insertion sites and to aid in elucidating the mode of insertion for vectors analyzed by LAM-PCR with Illumina MiSeq analysis off of the 3’ LTR, conventional PCR was performed off of the 5’ LTR. Forward genomic primers outside of the 5’ LTR were designed using the insertion site determined by 3’ LAM-PCR and subsequent sequencing. Reverse primers were designed within the 5’ LTR and conventional PCR was performed in order to amplify the 5’ vector-genome junction. The resulting PCR products were analyzed by Sanger sequencing.

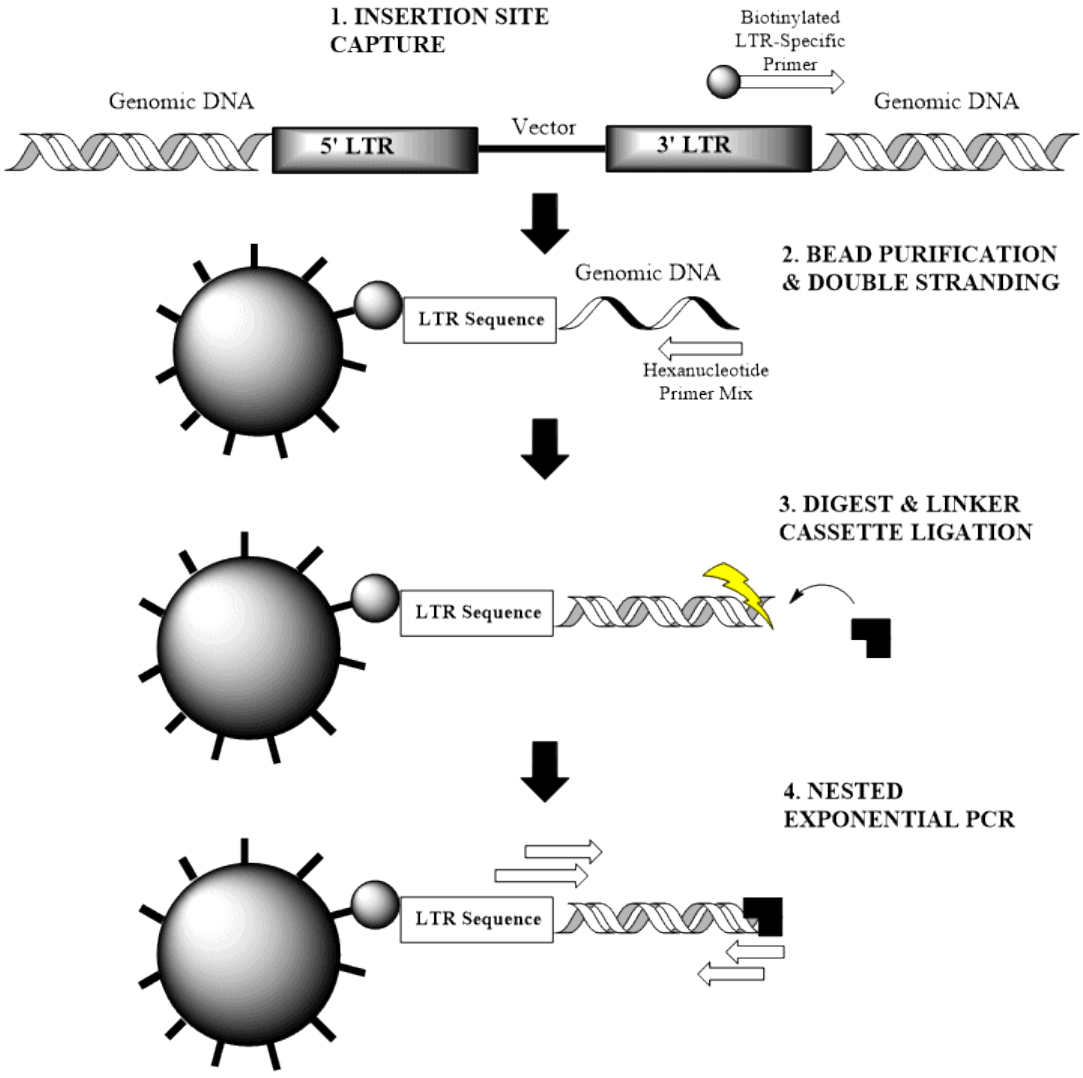
Sequence analysis

Pyrosequencing was performed on a Roche 454 FLX Titanium platform at the Indiana University Center for Genomics and Bioinformatics, Bloomington IN. 454 sequences were analyzed by Troy Hawkins, Ph.D. with the SeqMap 2.0 web server platform (Hawkins et al., 2011).

Sample libraries sequenced on an Illumina MiSeq platform were performed at the Genomics Core Facility of the University of Notre Dame. Paired-end reads of 250

base-pair were generated. Initial MiSeq analysis and in-house scripts were written by Hongyu Gao, Ph.D. In-house scripts were used to de-multiplex the reads and to retain paired-end reads that have both perfectly matched viral vector-specific primer and linker primer used in the last round of PCR. MiSeq adapter and linker sequences were trimmed off with Cutadapt version 1.5 (Martin, 2011). Paired-end reads were merged with PEAR version 0.9.5-64 (Zhang, Kobert, Flouri, & Stamatakis, 2014). For the non-overlapping paired-end reads, the reads containing the viral vector sequence were kept. The viral vector sequences in the merged reads were located with cross-match version .9909329 (http://www.phrap.org). Reads with less than 30 base-pair viral vector sequence and less than 10 base-pair remaining genomic sequence were discarded. The remaining reads were further clustered with usearch7.0.1090_i86linux32 (Edgar, 2010). Representative reads of each cluster were mapped to human reference genome hg19 using bwa version 0.7.5a (H. Li & Durbin, 2009). Alignments demonstrating 100% sequence identity were designated as integration sites. The cross-match output, cluster and mapping information were compiled together in R (RCore, 2012) and further analyzed.

Figure 18. Overview of LAM-PCR protocol. 1. Capture of integrated vector sequences by linear amplification off of 3' LTR. 2. Purification, selection and double stranding of captured vector-genome junctions. 3. Attaching a linker cassette to facilitate amplification of vector-genome junctions. 4. Nested PCR amplification to improve specificity of vector sequences and facilitate visualization and sequence analysis.



C. Results

Capture and analysis of NILV insertion sites by LAM-PCR lacks the sensitivity to account for features associated with non-integrase mediated integration

NILV offer a unique challenge for the capture and analysis of illegitimate non-integrase-mediated integrations. Existing PCR-based capture methods for vector insertion sites, such as LM-, LAM-, and nrLAM-PCR, utilize primers specific for vector LTR sequences to amplify the vector-genome junctions (Figure 18). The aberrant insertions sites observed with NILV include insertion and/or deletions at the vector genome junction along with truncations of the vectors LTRs, limiting the efficacy and sensitivity of PCR-based capture methods to analyze NILV insertion sites. We have evaluated the efficacy of analyzing NILV insertion sites by LAM-PCR. An initial experiment was performed with sequencing of LAM-PCR products by Roche 454 pyrosequencing and a second more in depth experiment on an Illumina MiSeq platform with subsequent confirmation of insertion sites by Sanger sequencing of the 5' vector-genome junction.

The integration competent lentiviral vector (CS-CZW) sequenced by MiSeq demonstrated the features associated with integrase-mediated insertion (Table 2). Vector insertion sites were identified for 10 of 12 clones at the 3' vector-genome junction. Analysis of the 5' vector-genome junction was verified for 6 randomly selected clones, confirming the insertion sites. The insertions sites displayed the terminal CA dinucleotide, indicative of integrase-mediated end-processing, and the five base-pair flanking repeat at the vector-genome junctions.

Analysis of integrase-defective NILV have demonstrated non-integrase-mediated mechanisms of integration such as non-homologous end-joining and homologous recombination (Koyama et al., 2013; Matrai et al., 2011; Nightingale et al., 2006). The resulting insertion and/or deletions at the vector-genome junction can inhibit the binding of primers for vector capture. Our initial attempt at analyzing the integrase mutant (IN/D116N) by LAM-PCR on the 454 platform allowed for the capture of 57% (4 of 7 clones) at the 3' vector-genome junction (Table 3). All of the captured insertion sites demonstrated LTR truncations and insertions/deletions at the junction. To verify our results, we repeated this analysis with more clones utilizing the MiSeq platform and further verification of insertion sites by conventional PCR and Sanger sequencing of the 5' vector-genome junction in vectors with confirmed 3' insertion sites. With confirmation of 58% (7 of 12 clones) of the insertion sites, there was no improvement in capture observed with MiSeq as compared to 454 using this method. Of the confirmed insertion sites 57% (4 of 7 clones) displayed the aberrant insertion sites associated with NILV integrations. 2 out of the 7 clones with confirmed 3' insertion sites were confirmed by 5' analysis. Both of the clones displayed the aberrant features associated non-integrase-mediated integration, providing confidence to our results observed by 454 sequencing. However, only confirming 28% of the 3' insertion sites demonstrates the lack of sensitivity to detect NILV insertions by PCR-based methods relying on LTR-specific primers.

Table 2. Normal integrating lentiviral vector insertion site analysis demonstrates canonical integrase-mediated insertion site features. HEK 293 clones transduced with the normal integrating lentiviral vector (CS-CZW) analyzed by LAM-PCR with sequencing by Illumina MiSeq. [a] indicates whether clones analyzed map to a reference genome; [b] indicates any deletions to the terminal ends of the LTR; [c] indicates if vectors with 3' confirmation had 5' vector-genome junction analysis performed; [d] indicates the presence of a 5 base-pair flanking repeat at the vector-genome junctions of vectors with 5' and 3' confirmation. Empty cells indicate the clone was not selected for 5' analysis.

Clone	3' IS Identified[a]	3' LTR Truncations[b]	5' Confirmation Performed?[c]	5' IS Identified[a]	5' LTR Truncations[b]	5-bp Repeat[d]
m1	Yes	No	Yes	Yes	No	Yes
m2	No	No	No			
m3	Yes	No	No			
m4	Yes	No	Yes	Yes	No	Yes
m5	Yes	No	Yes	Yes	No	Yes
m6	Yes	No	No			
m7	Yes	No	Yes	Yes	No	Yes
m8	Yes	No	No			
m9	Yes	No	Yes	Yes	No	Yes
m10	No	No	No			
m11	Yes	No	Yes	Yes	No	Yes
m12	Yes	No	No			

Table 3. Integrase defective mutant NILV insertion site analysis demonstrates features associated with non-integrase-mediated insertion. HEK 293 clones transduced with the integrase defective NILV (IN/D116N) analyzed by LAM-PCR with sequencing by Illumina MiSeq. [a] indicates whether clones analyzed map to a reference genome; [b] indicates any deletions to the terminal ends of the LTR; [c] indicates if vectors with 3' confirmation had 5' vector-genome junction analysis performed; [d] indicates the presence of genomic deletions or insertions at the vector-genome junctions; [e] indicates the presence of a 5 base-pair flanking repeat at the vector-genome junctions of vectors with 5' and 3' confirmation. Empty cells indicate 3' or 5' insertion site was not identified and additional analysis was not able to be performed.

Clone	Sequencing Platform	3' IS Identified[a]	3' LTR Truncations[b]	5' Confirmation Performed?[c]	5' IS Identified[a]	5' LTR Truncations[b]	Junction INDELs[d]	5-bp Repeat[e]
r1	Roche 454	No						
r2	Roche 454	No						
r3	Roche 454	Yes	Yes	No			Yes	
r4	Roche 454	Yes	Yes	No			Yes	
r5	Roche 454	No						
r6	Roche 454	Yes	Yes	No			Yes	
r7	Roche 454	Yes	Yes	No			Yes	
m13	Illumina MiSeq	No						
m14	Illumina MiSeq	No						
m15	Illumina MiSeq	Yes	Yes	Yes	Yes	Yes	Yes	No
m16	Illumina MiSeq	Yes	Yes	Yes	Yes	No	Yes	No
m17	Illumina MiSeq	No						
m18	Illumina MiSeq	Yes	No	Yes	No		Yes	
m19	Illumina MiSeq	Yes	No	Yes	No		No	
m20	Illumina MiSeq	Yes	No	Yes	No		No	
m21	Illumina MiSeq	Yes	No	Yes	No		No	
m22	Illumina MiSeq	Yes	Yes	Yes	No		No	
m23	Illumina MiSeq	No						
m24	Illumina MiSeq	No						

The Roche 454 sequence analysis of our novel LTRΔatt NILV was able to identify insertion sites in 50% (4 of 8) of clones using 3' LAM-PCR (Table 4). The results revealed intact 3' LTRs in 75% (3 of 4 clones) of these. However, 75% of these 4 verified 3' insertion sites presented insertions at the 5' vector genome junction ranging from 5-168 base-pairs. The MiSeq analysis of LTRΔatt clones was able to confirm 73% of the 3' insertions sites. Only one of the 3' confirmed insertions sites revealed a truncation of the LTR and none presented insertions or deletions at the vector-genome junction. Only two of the 3' confirmed insertions sites were able to be further confirmed by 5' analysis. These two clones both presented insertions and deletions at the 5' vector-genome junction and one of them indicated the presence of a flanking 5 base-pair repeat of genomic DNA. These results indicated that the LTRΔatt vector was demonstrating features of integrase-mediated insertion at the 3' vector-genome junction while presenting aberrant insertion at the 5' LTR junction.

Table 4. U3 LTR integrase DNA attachment site deleted NILV insertion site analysis demonstrates features associated with both integrase- and non-integrase-mediated insertion. HEK 293 clones transduced with the U3 attachment site deleted NILV (LTRΔatt) analyzed by LAM-PCR with sequencing by Illumina MiSeq. [a] indicates whether clones analyzed map to a reference genome; [b] indicates any deletions to the terminal ends of the LTR; [c] indicates if vectors with 3' confirmation had 5' vector-genome junction analysis performed; [d] indicates the presence of genomic deletions or insertions at the vector-genome junctions; [e] indicates the presence of a 5 base-pair flanking repeat at the vector-genome junctions of vectors with 5' and 3' confirmation; * insertions and deletions were only present at the 5' LTR vector-genome junction. Empty cells indicate 3' or 5' insertion site was not identified and additional analysis was not able to be performed.

Clone	Sequencing Platform	3' IS identified[a]	3' LTR Truncations[b]	5' Confirmation Performed?[c]	5' IS Identified[a]	5' LTR Truncations[b]	Junction INDELs[d]	5-bp Repeat[e]
r8	Roche 454	No						
r9	Roche 454	Yes	No	No			Yes	
r10	Roche 454	Yes	No	No			No	
r11	Roche 454	No						
r12	Roche 454	Yes	Yes	No			Yes	
r13	Roche 454	Yes	No	No			Yes	
r14	Roche 454	No						
r15	Roche 454	No						
m25	Illumina MiSeq	Yes	No	Yes			No	
m26	Illumina MiSeq	Yes	No	Yes			No	
m27	Illumina MiSeq	Yes	No	Yes			No	
m28	Illumina MiSeq	No						
m29	Illumina MiSeq	Yes	No	Yes			No	
m30	Illumina MiSeq	Yes	No	Yes	Yes	Yes	Yes*	Yes
m31	Illumina MiSeq	Yes	Yes	Yes			No	
m32	Illumina MiSeq	No						
m33	Illumina MiSeq	Yes	No	Yes			No	
m34	Illumina MiSeq	Yes	No	Yes	Yes	Yes	Yes*	No
m35	Illumina MiSeq	No						

We have previously demonstrated that combining the LTRΔatt and IN/D116N mutations was able to significantly reduce the frequency of illegitimate integration as compared to either of these mutations independently (Figure 11). Our conclusion elucidated from the results in Table 4 indicate that the LTRΔatt mutation in the presence of a functional integrase protein can still bind to integrase at the intact U5 attachment site and facilitate integrase mediated insertion of the vectors 5' LTR. The double mutant vector would combine the attachment site deletion with an integrase protein defective for strand transfer and integration. Our sequencing of LAM-PCR results by MiSeq and 454 sequencing demonstrate the increased efficacy of inhibiting integration by the LTRΔatt-IN/D116N combination. We were able to confirm two of four 3' insertion sites by 454 sequencing, with both clones demonstrating the features associated with non-integrase-mediated insertion. Our MiSeq results for this vector were only able to identify vector insertion sites in 22% (2 of 9) of the analyzed clones for 3' insertion sites and we were unable to confirm any of the corresponding 5' insertion sites. The confirmed vectors displayed completely intact LTRs without the two base-pair truncation indicative of integrase-mediated end-processing. These limited results demonstrate that the inclusion of the integrase mutation in combination with the LTRΔatt mutation results in features similar to the IN/D116N vector, indicating the inhibition of features associated with integrase-mediated insertion at the 3' vector-genome junction as observed with LTRΔatt. They also demonstrate a lack of sensitivity to significantly analyze NILV clones using PCR-based techniques relying on intact LTR sequences for vector capture.

Table 5. Double mutant NILV insertion site analysis demonstrates features associated with non-integrase-mediated insertion. HEK 293 clones transduced with the double mutant NILV (LTRΔatt-IN/D116N) analyzed by LAM-PCR with sequencing by Illumina MiSeq. [a] indicates whether clones analyzed map to a reference genome; [b] indicates any deletions to the terminal ends of the LTR; [c] indicates if vectors with 3' confirmation had 5' vector-genome junction analysis performed; [d] indicates the presence of genomic deletions or insertions at the vector-genome junctions; [e] indicates the presence of a 5 base-pair flanking repeat at the vector-genome junctions of vectors with 5' and 3' confirmation. * 2 base-pair insertions indicating a lack of end-processing mediated by the integrase protein. Empty cells indicate 3' or 5' insertion site was not identified and additional analysis was not able to be performed.

Clone	Sequencing Platform	3' IS Identified[a]	3' LTR Truncations[b]	5' Confirmation Performed?[c]	5' IS Identified[a]	5' LTR Truncations[b]	Junction INDELs[d]	5-bp Repeat[e]
r16	Roche 454	Yes	Yes	No			Yes	
r17	Roche 454	No						
r18	Roche 454	Yes	Yes	No			Yes	
r19	Roche 454	No						
m36	Illumina MiSeq	No						
m37	Illumina MiSeq	No						
m38	Illumina MiSeq	No						
m39	Illumina MiSeq	No						
m40	Illumina MiSeq	No						
m41	Illumina MiSeq	No						
m42	Illumina MiSeq	Yes	No	Yes	No		Yes*	
m43	Illumina MiSeq	Yes	No	Yes	No		Yes*	
m44	Illumina MiSeq	No						

Modified insertion site analysis improves the sensitivity for capturing non-integrase-mediated integrations by NILV

The lack of sensitivity observed when using LAM-PCR for the capture and analysis of aberrant insertions sites associated with NILV led us to explore other methodologies which would provide more flexibility for improving our analysis. Recent reports have demonstrated a method utilizing sonication of transduced DNA in order to produce fragments of a manageable size which would allow for ready manipulation and amplification (De Ravin et al., 2014; S. Zhou et al., 2014). We sought to modify this methodology to improve our analysis by accounting for the high frequency of truncations to NILV LTRs (Figure 19). (1) We began with a sonication program for the production of larger fragments to allow for capture and further purification of integrated vector sequences. (2) Linker cassettes were ligated onto the ends of the fragmented genome. (3) We began our initial capture with primers set outside of the LTR in the vector sequence to compensate for the occurrence of deletions. Nested PCR was performed to increase specificity and to keep the amplicons within a size range which would allow for complete coverage by paired-end sequencing. These modifications permit up to approximately 280 base-pair LTR deletions without affecting sensitivity. Sequence results of the IN/D116N integrase mutant supported our previous findings and those reported by others (Table 6). Sequencing produced an average of 171K reads per clone. The modified methodology was able to improve the capture of insertions sites from ~58% by LAM-PCR to 83%. The majority of the insertion sites demonstrated large LTR truncations that would have inhibited capture by LAM-PCR.

Figure 19. Modified insertion site analysis methodology. Insertion site analysis of selected vector transduced clones intended to improve the capture of NILV insertion sites as described in the text.

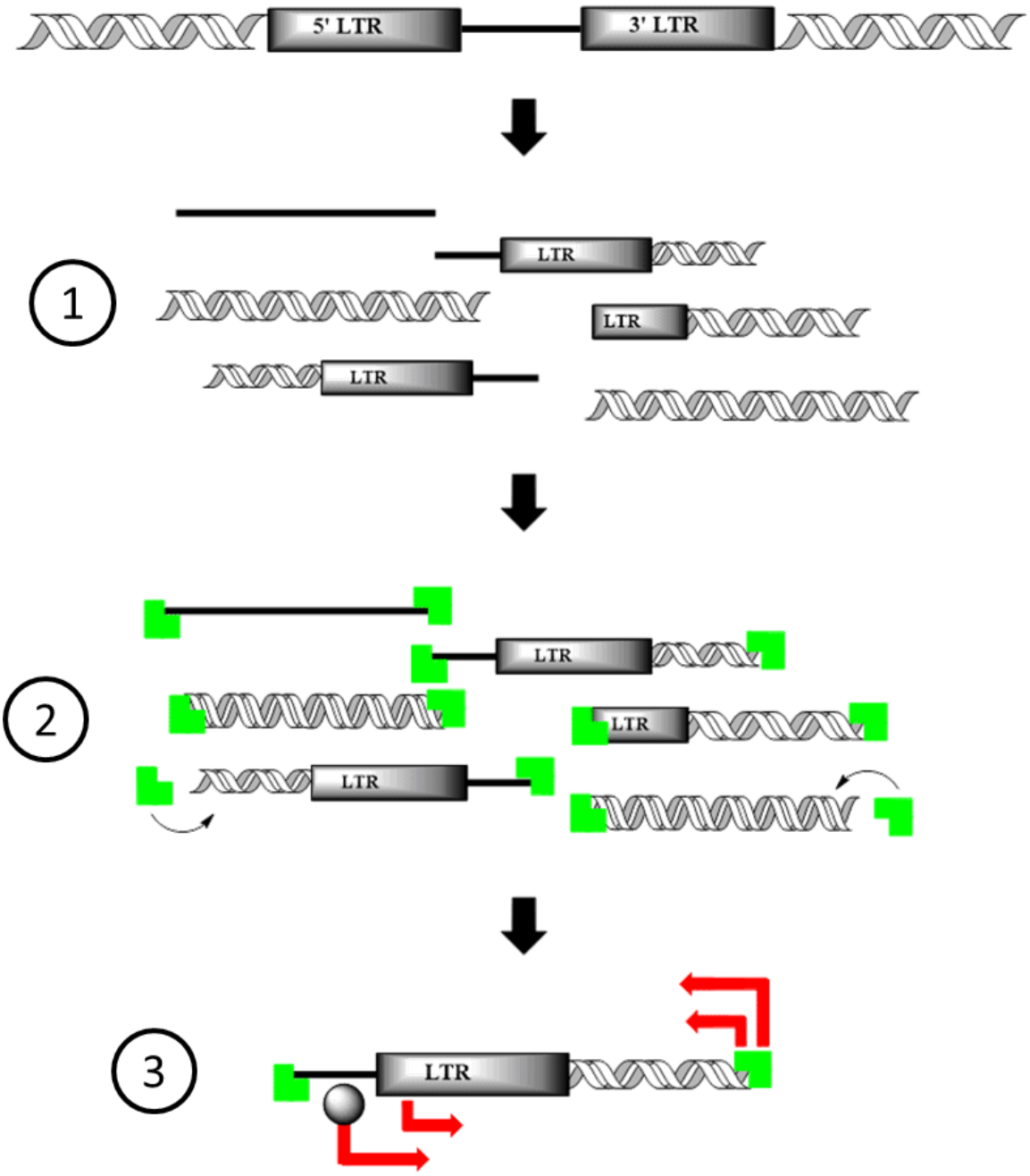


Table 6. Modified insertion site methodology demonstrates improved sensitivity for the analysis of Integrase defective NILV integrations. HEK 293 clones transduced with an integrase defective NILV (IN/D116N) analyzed by a modified vector insertion site methodology with sequencing by Illumina MiSeq. [a] indicates reads map to a reference genome meeting filtering criteria; [b] size of deletions (base-pairs) to the terminal ends of the LTR; [c] length of insertions in base-pairs detected at the vector-genome junction; [d] indicates the 3' and 5' confirmed reads map to the same genomic location; [e] indicates the presence of any genomic DNA deletions at the vector-genome junction; [f] indicates the presence of a 5 base-pair flanking repeat at the vector-genome junctions. Empty cells indicate 3' or 5' insertion site was not identified and additional analysis was not able to be performed.

Clone	3' IS Ident.[a]	3' LTR Trunc.[b]	3' Insert.[c]	5' IS Ident.[a]	5' LTR Trunc.[b]	5' Insert.[c]	3' & 5' Paired[d]	Genomic Deln.[e]	5-bp Repeat[f]
c1	Yes	2	0	Yes	0	0	Yes	Yes	No
c2	Yes	125	4	Yes	215	0	No		
c3	Yes	0	3	Yes	241	4	No		
c4	No			No			No		
c5	Yes	138	42	Yes	235	41	No		
c6	Yes	170	36	Yes	226	45	No		
c7	Yes	0	8	Yes	0	0	Yes	Yes	No
c8	Yes	0	0	Yes	10	0	Yes	Yes	No
c9	Yes	8	0	Yes	11	0	Yes	Yes	No
c10	Yes	170	40	Yes	111	83	No		
c11	No			No			No		
c12	Yes	9	0	Yes	159	98	No		

Sequencing results from modified insertion site analysis of the novel U3 LTR attachment site deleted NILV (LTRΔatt) also support our LAM-PCR results (Table 7). The results strongly support our hypothesis that a functional integrase protein is able to bind to the intact U5 attachment site and facilitate integration at the 3' LTR, while ablation of the U3 attachment site results in aberrant insertion at the 5' LTR. This was evident as only 11% (1 of 9) of the confirmed 3' insertion sites presents a truncation of the LTR, while 100% of the confirmed 5' insertion sites demonstrate the large deletions and insertions associated with non-integrase-mediated integration. However, while these results are promising, there was no improvement in the percentage of clones confirmed by this method. Further optimization of this methodology and a larger sample size could improve the capture of vector insertion sites.

Table 7. Modified insertion site methodology supports PCR-based insertion site capture results, identifying a unique mode of integration by a novel U3 LTR integrase attachment site mutant NILV. HEK 293 clones transduced with a novel U3 LTR attachment deleted NILV (LTRΔatt) analyzed by a modified vector insertion site methodology. [a] indicates reads map to a reference genome meeting filtering criteria; [b] size of deletions (base-pairs) to the terminal ends of the LTR; [c] length of insertions in base-pairs detected at the vector-genome junction; [d] indicates the 3' and 5' confirmed reads map to the same genomic location; [e] indicates the presence of any genomic DNA deletions at the vector-genome junction; [f] indicates the presence of a 5 base-pair flanking repeat at the vector-genome junctions. Empty cells indicate 3' or 5' insertion site was not identified and additional analysis was not able to be performed.

Clone	3' IS Ident.[a]	3' LTR Trunc.[b]	3' Insert.[c]	5' IS Ident.[a]	5' LTR Trunc.[b]	5' Insert.[c]	3' & 5' Paired[d]	Genomic Deln.[e]	5-bp Repeat[f]
c13	Yes	212	0	Yes	242	72	Yes	Yes	No
c14	Yes	0	0	Yes	208	0	No		
c15	Yes	0	0	Yes	225	14	No		
c16	Yes	0	0	No			No		
c17	No			No			No		
c18	No			No			No		
c19	Yes	0	0	Yes	242	68	No		
c20	No			No			No		
c21	No			No			No		
c22	No			No			No		
c23	No			No			No		
c24	No			No			No		
c25	No			Yes	229	0	No		
c26	Yes	0	0	No			No		
c27	Yes	0	0	No			No		
c28	Yes	0	0	Yes	233	0	No		
c29	Yes	0	0	Yes	193	4	No		
c30	No			Yes	232	0	No		

Our previous results when analyzing the insertion sites of the double mutant vector (LTRΔatt-IN/D116N) yielded a very low frequency of clones we were able to confirm. It was likely that the inclusion of both mutations for inhibiting integration resulted in a higher frequency of aberrant insertion sites with large LTR truncations which may have prevented the capture by LAM-PCR. The modified insertion site analysis resulted in a significantly improved yield of LTRΔatt-IN/D116N clones we were able to confirm. Similar to the LTRΔatt insertion sites, the double mutant presented aberrant 5' vector-genome junctions in 93% of confirmed clones. However, unlike the LTRΔatt mutant, the 3' LTR junction demonstrated aberrations in 100% of the clones analyzed. Though, 50% of these presented with insertions and no truncations to the vectors LTR.

Table 8. Double mutant NILV demonstrate increased frequencies of aberrations at vector insertions sites relative to independent mutations. HEK 293 clones transduced with a double mutant NILV (LTRΔatt-IN/D116N) analyzed by a modified vector insertion site methodology with sequencing by Illumina MiSeq. [a] indicates reads map to a reference genome meeting filtering criteria; [b] size of deletions (base-pairs) to the terminal ends of the LTR; [c] length of insertions in base-pairs detected at the vector-genome junction; [d] indicates the 3' and 5' confirmed reads map to the same genomic location; [e] indicates the presence of any genomic DNA deletions at the vector-genome junction; [f] indicates the presence of a 5 base-pair flanking repeat at the vector-genome junctions; * analysis detected an aberrant 80 base-pair repeat flanking the vector.

Clone	3' IS Ident.[a]	3' LTR Trunc.[b]	3' Insert.[c]	5' IS Ident.[a]	5' LTR Trunc.[b]	5' Insert.[c]	3' & 5' Paired[d]	Genomic Deln.[e]	5-bp Repeat[f]
c31	Yes	5	5	Yes	0	0	Yes	No	No*
c32	Yes	196	43	Yes	170	0	No		
c33	Yes	191	0	Yes	232	0	No		
c34	Yes	0	74	Yes	215	139	No		
c35	No			No			No		
c36	Yes	1	3	Yes	214	0	No		
c37	No			No			No		
c38	No			No			No		
c39	Yes	0	18	Yes	231	0	No		
c40	No			No			No		
c41	No			Yes	222	0	No		
c42	No			No			No		
c43	Yes	0	14	Yes	215	201	No		
c44	Yes	0	18	Yes	215	49	No		
c45	Yes	2	2	Yes	161	7	Yes	Yes	No
c46	Yes	0	32	Yes	163	2	No		
c47	Yes	0	7	Yes	0	18	Yes	No	No
c48	No			No			No		
c49	Yes	4	12	Yes	229	1	No		
c50	No			No			No		
c51	No			No			No		
c52	No			Yes	229	15	No		

D. Discussion

The studies presented here demonstrate the necessity for the optimization of methodologies for the analysis of illegitimate integrations by NILV. The novel LTRΔatt vector has demonstrated a significant reduction in the frequency of illegitimate integration when combined with an integrase catalytic core mutation (IN/D116N) as compared to either of these independently (Figure 11). We sought to characterize the insertion sites of clones transduced with NILV containing these mutations in order to discern the mechanisms of integration leading to the variability observed in integration between them and the improved profile when combined. Existing methodologies for analyzing vector insertion sites rely upon capture and amplification using primers at the vector-genome junction. As we have demonstrated here, this presents limitations when analyzing NILV which integrate by non-integrase-mediated mechanisms, such as NHEJ and HR, resulting in LTR truncations and insertions/deletions at the vector-genome junction. We have demonstrated that modifying insertion site capture methods to account for these variations can improve the analysis of NILV.

Our results demonstrate the canonical features associated with the types of vectors used (Figure 20). As was expected, integrase-mediated insertion was prominent when using an integration competent vector. We have shown our integrase defective NILV to display illegitimate integration features indicative of non-integrase-mediated insertion, similar to reports by others when using integrase catalytic core mutants (Matrai et al., 2011; Nightingale et al., 2006).

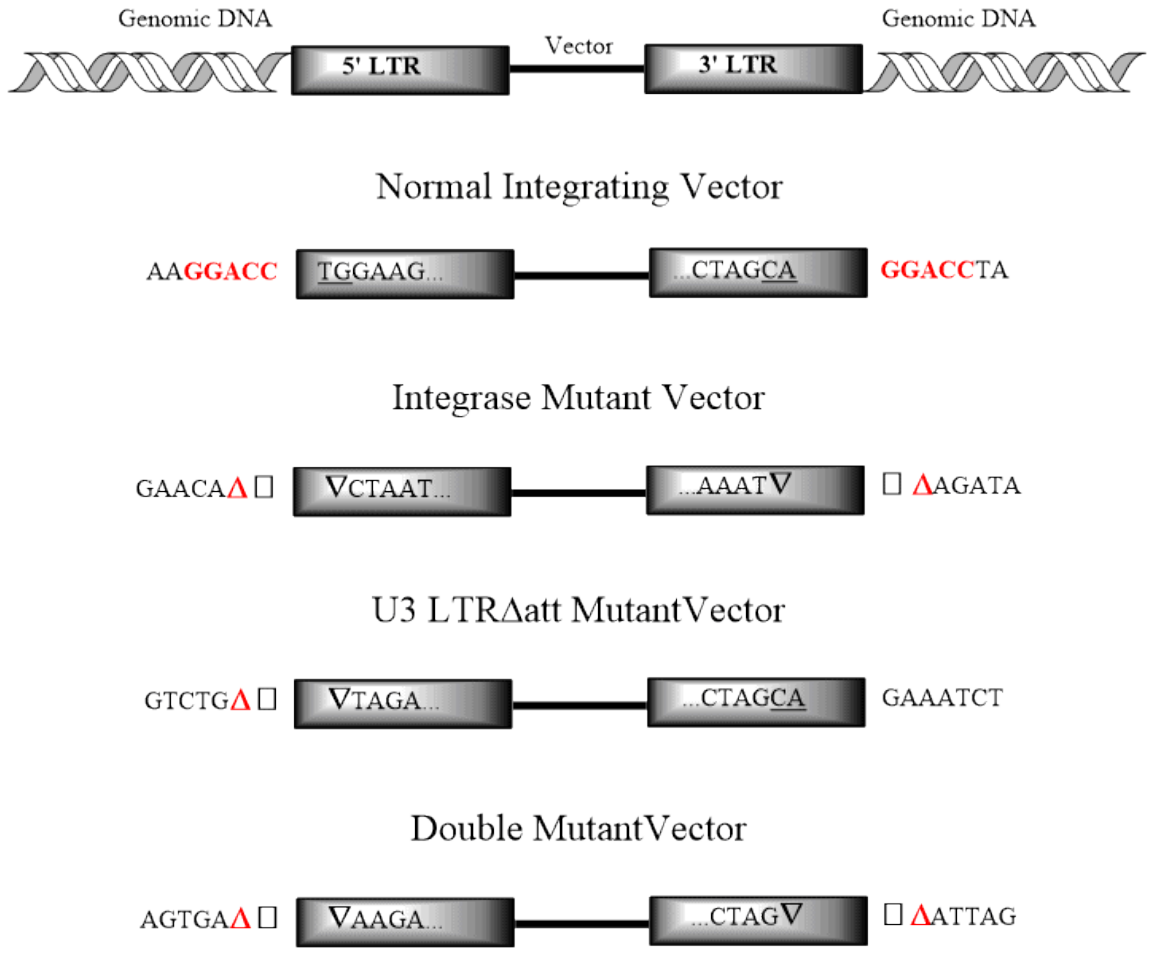
Our novel LTRΔatt vector demonstrates a unique mode of illegitimate integration at one end of the vector. The deletion of the U3 attachment site at the terminal end of the 5' LTR appears to effectively inhibit integrase binding and insertion. However, our results suggest that integrase can still bind with the intact U5 attachment site at the 3' LTR. Integrase binding appears to stabilize this terminal end of the vector and/or facilitate integration into the host genome, indicated by the lack of 3' LTR truncations observed and indications of integrase-mediated end-processing (Table 7).

As combining the attachment site and integrase mutations leads to a significant reduction in integration frequency, we expect the defective integrase protein would inhibit the features associated with strand-transfer at the 3' LTR. The combined LTRΔatt-IN/D116N NILV demonstrates similar features to LTRΔatt at the 5' LTR, with large LTR truncations and insertions/deletions at the vector genome junction. Inclusion of the integrase mutation resulted in the presence of features at the 3' LTR indicative of non-integrase mediated insertion, similar to the IN/D116N mutation independently. Interestingly, it was observed that insertions at the 5' LTR of clones transduced with vector containing the ablated U3 attachment site demonstrated significantly more and larger LTR truncations than were observed for the 3' LTR with its intact U5 attachment site regardless of integrase inhibition. Common mutations to produce integrase-defective NILV introduce missense mutations to a triad of residues at the catalytic core of integrase that inhibit strand-transfer (Kulkosky et al., 1992; Shibagaki & Chow, 1997). These mutations do not inhibit the ability of integrase to bind vector DNA, which is controlled by a Q148 residue in the core domain of integrase (Esposito & Craigie, 1998;

Gerton, Ohgi, Olsen, DeRisi, & Brown, 1998). Thus, it appears that in the presence of intact LTR integrase attachment sites, integrase mutants retain the ability to bind vector DNA stabilizing the terminal ends and inhibiting degradation resulting in fewer truncations of vector LTR sequences.

In summary, these results demonstrate that improving methods to analyze the aberrant insertion sites of NILV will require capture and amplification strategies which can compensate for insertions and deletions at the vector-genome junctions. With further optimization our modified in insertion site analysis method should enable the capture of the majority of NILV insertions as may be required for FDA approval to get NILV into the clinic. Our findings illustrate the characteristics of our novel U3 LTR attachment site mutation for inhibiting integration and how combination with integrase mutations can improve its inhibition of integrase-mediated insertion. This demonstrates the feasibility of combining other mutations to further improve the utility of NILV.

Figure 20. Overview of vector-genome junction features of vector transduced clones. Summary of our results obtained from sequencing of Zeocin selected clones transduced with a normal integrating vector (CS-CZW), an integrase catatlytic core mutant NILV (IN/D116N), a U3 LTR integrase attachment sited deleted NILV (LTRΔatt) and a double mutant (LTRΔatt-IN/D116N). Red highlighted base-pairs indicate flanking 5-bp repeat of genomic DNA by integrase-mediated strand-transfer; underlined TG/CA dinucleotides represent end-processing by a function integrase protein; inverted black triangles represent LTR truncations; red highlighted triangle represents deletions of genomic DNA; black squares indicate insertions at the vector-genome junction.



V. Conclusions

The use of viral vectors for the gene therapy treatment of patients poses inherent risks associated with biological tools discovered by their ability to induce disease. Early clinical trials utilizing gamma retroviral vectors have resulted in adverse events to patients induced by insertional mutagenesis, highlighting these risks. Lentiviral vector mediated gene therapy has been explored as a safer alternative to early gamma retroviral vectors and have demonstrated an improved profile relative to the risk of insertional mutagenesis. As these vectors have begun to enter various clinical trials it is imperative to ensure that we provide the safest product possible when treating patients. The necessity for improving vector safety ranges from the production of vector supernatant, the design and safety profile of the vectors used for transduction, to follow up analysis of vector transduced cells. We have evaluated an FDA approved DNase for improving the purity of lentiviral vector supernatant, developed a novel NILV to improve the utility of these vectors for clinical use, and modified insertion site analysis methods to improve the capture of aberrant insertion sites associated with illegitimate NILV integration.

Lentiviral vector production utilizes a transient transfection of producer cell lines, such as HEK 293T, with a four plasmid system. This system places the vector components in trans to each other in order to reduce the risk of recombination and the production of a replication competent vector. In order to further reduce the risk of recombination and the exposure of patients to viral DNA, residual plasmid DNA must be

thoroughly removed following transfection. Here we have demonstrated the application of Pulmozyme, a recombinant human DNase FDA approved for treating cystic fibrosis patients (Pan et al., 2001; Shak et al., 1990; Shire, 1996; Ulmer et al., 1996). Pulmozyme was investigated as an alternative to Benzonase, a bacterially derived DNase. While Pulmozyme was determined to be an effective, clinically relevant alternative to Benzonase, the possibility of residual DNase in the vector supernatant could pose a risk to patients. Adverse events due to Pulmozyme are rare and it has been shown to rapidly degrade in serum with a half-life of three to four hours (Roche, 2014). However, gene therapy of immune-compromised individuals would require increased purity of vector supernatant to prevent adverse effects from residual production components. In order to produce clinical grade vector supernatant it will be necessary to implement downstream processing to remove impurities and to include detection of residual DNase along with the rigorous testing already mandated prior to clinical application.

Development of the novel LTRΔatt NILV demonstrated reductions in integration frequency inferior to integrase defective NILV. However, when combined, these two mutations significantly reduced the frequency of integration relative to either independently. Combining these two mutations with a 3' PPT deletion to further inhibit integration and a large U3 deletion of cis-acting inhibitors to transgene expression resulted in an improvement in transgene expression levels while significantly reducing the frequency of illegitimate integration. This provided an important proof of concept for improvement of NILV to increase their clinical utility.

In order to further characterize the novel LTRΔatt vector, we sought to sequence the vector insertion sites in order to elucidate the mechanism resulting in the increased illegitimate integration relative to the IN/D116N integrase defective NILV. However, as we and others have demonstrated (Gaur & Leavitt, 1998; Matrai et al., 2011; Nightingale et al., 2006), NILV insertion sites demonstrate aberrant vector-genome junctions which can inhibit their capture by PCR-based methods utilizing primers within the LTRs for amplification. We sought to modify a method of vector insertion site capture utilizing sonication of the transduced vector genome (De Ravin et al., 2014; S. Zhou et al., 2014) prior to initial vector capture. Implementing this new methodology along with our modifications, significantly improved our ability to capture and analyze NILV insertion sites. Further optimization of this methodology will allow for even more improvement in the capture of NILV insertion sites for analysis. Possible modifications include increasing the specificity to reduce the capture of non-specific sequences and improved capture of vector-containing fragments. To improve the capture of NILV insertion sites it may be possible to devise a pull down to capture or bind all fragments containing vector sequences. These fragments could then be amplified indiscriminately utilizing terminal linker cassette/adapter sequences, thus eliminating the necessity to rely upon primers near the variable vector-genome junctions of NILV. This methodology would require an increased bioinformatics analysis in order to filter and analyze reads containing vector-genome junctions. These optimizations could theoretically improve our ability to analyze aberrant NILV insertion sites increasing the clinical utility of these vectors by allowing for more sensitive analysis of transduced cells post-transduction.

Taken in aggregate, our sequencing results demonstrate the expected features for our normal integrating vector (CS-CZW) and the integrase-defective mutant NILV (IN/D116N), in line with previous reports. Our novel LTRΔatt NILV demonstrated features unique to this particular mutant. Our results indicate that deletion of the U3 LTR integrase attachment site effectively inhibits integrase protein binding at the terminal end of the vectors 5' LTR. Sequencing of the 3' insertion sites indicate that the presence of an intact U5 attachment site at the vectors 3' LTR allows for integrase binding and insertion at this end. The 5' end is then subsequently integrated likely by non-homologous end-joining resulting in the large truncations at the terminal end of the LTR and insertions/deletions of the adjacent genomic DNA. Combining the LTRΔatt and IN/D116N mutations confirms these findings. The presence of the integrase mutation results in an increased frequency of aberrant features at the 3' LTR vector-genome junction similar to those observed with the IN/D116N NILV independently. However, a significantly lower frequency of LTR truncations in the presence of only the IN/D116N mutation indicate that integrase is still able to bind the intact attachment sites and stabilize the ends from degradation prior to non-integrase-mediated insertion. These results indicate that including functional mutations to the U5 attachment site along with the LTRΔatt and IN/D116N mutations may further reduce the frequency of integration by NILV. As complete ablation of the U5 attachment site does not result in viable vectors, the inclusion of point mutations, similar to those previously reported (Apolonia et al., 2007; Nightingale et al., 2006), could provide improved inhibition of integration.

These results demonstrate how elucidating the features associated with illegitimate integration by NILV can improve their design for use in the clinic.

Possible improvements to this line of work include the development of assays to evaluate the presence of residual DNase enzymes in the final vector supernatant product for clinical applications. An enzyme-linked immunosorbent assay, with antibodies to DNase I, could provide for a method to satisfy this necessity.

As described above, further optimization of methods is necessary to improve the capture of aberrant insertion sites associated with NILV integrations. Improving the sensitivity to analyze NILV insertion sites will increase the utility of these vectors for use in the clinic.

Sequencing and integration analysis results demonstrate the utility of combining modifications to reduce the frequency of illegitimate integration by NILV. Sequence results demonstrating the absence of any integrase-mediated integration features at LTRs with ablated attachment sites present the possibility of including an additional mutation to the U5 LTR integrase attachment site to further reduce the frequency of integration by NILV.

To further improve NILV utility these modifications could be combined with inhibition of cellular restriction factors to transgene expression which have been shown to improve NILV episomal transgene expression to levels associated with integration competent vectors (Berger et al., 2009; Negri et al., 2012; Pelascini et al., 2013). The production of an NILV with optimal transgene expression would provide for a safer vector for clinical use by decreasing the volume of vector necessary to provide a

therapeutic benefit. Decreasing the quantity of vector required for gene therapy applications translates to a decreased risk of illegitimate integrations leading to insertional mutagenesis. Combining this with multiple modifications for effectively inhibiting integrase-mediated insertion would result in an optimal NILV for use.

In conclusion, this study highlights how continuous improvements can advance the safety of lentiviral vector mediated gene therapy. Reducing the impurities in vector supernatant, providing the safest vector possible, and improving the sensitivity for analyzing vector insertion sites post-treatment will help to move lentiviral vectors forward as they continue to enter the clinic.

VI. Future Direction

"Adapted from Shaw A, Cornetta K. Design and Potential of Non-Integrating Lentiviral Vectors. Biomedicines. 2014; 2(1):14-35"

We have demonstrated significant improvements in expression and the reduction of illegitimate integration when combining modifications to NILV design. We have shown that combining a novel U3 LTR integrase DNA attachment site deletion with an in integrase mutation (D116N), a deletion of the 3' PPT, and a large deletion of the U3 region demonstrates improved transgene expression and reduced illegitimate integration as compared to current NILV. This novel combination of NILV mutations provides an optimal vector for many clinical applications.

There are many clinical applications which could benefit from our optimized NILV with its improved safety profile. These applications include vaccinations, cell-type and lineage differentiation, as donor templates for homologous recombination in site-directed integration systems, and as delivery systems for cytotoxic cancer therapies. NILV are also being considered for gene transfer into slowly growing or non-dividing tissues where persistent episomal expression can provide a long-lasting therapeutic effect.

Vaccination is an application where only transient expression is required and NILVs have been shown to stimulate an efficient and sustained immune response (Daenthanasanmak et al., 2012; Hu, Dai, & Wang, 2010; Michelini et al., 2009; Negri et al., 2012). Preclinical studies of NILVs have demonstrated immune responses against

human papillomavirus (HPV), malaria, HIV-1 and the hepatitis B and C viruses (Coutant et al., 2008; Coutant et al., 2012; Deng et al., 2013; Grasso et al., 2012; Karwacz et al., 2009; Negri et al., 2007) thus showing their potential for use in vaccine development.

A second application where transient expression is preferred over sustained expression is in cell reprogramming. This includes creation of induced pluripotent stem (iPS) cells and differentiation of iPS or embryonic stem (ES) cells into a lineage of interest. While somatic cells have been successfully reprogrammed into iPS cells using integrating vectors (Papapetrou et al., 2009), the factors for inducing pluripotency are not necessary beyond initial reprogramming and constitutive expression of the factors has been shown to be harmful (Okita, Ichisaka, & Yamanaka, 2007; Papapetrou et al., 2009; Takahashi et al., 2007; Wernig et al., 2007). Continued expression can be oncogenic (Okita et al., 2007) and can also affect differentiation of iPS cells into other lineages (Soldner et al., 2009). Transgene free iPS cells have been produced using integrating LV followed by excision with NILVs expressing Cre recombinase after reprogramming (Papapetrou & Sadelain, 2011) as well as by transient expression using both non-integrating adenoviral vectors (Stadtfeld, Nagaya, Utikal, Weir, & Hochedlinger, 2008; Yu et al., 2009) and NILV (Mali et al., 2008). NILV have also been successful in differentiating ES cells into specific progenitors (Yang et al., 2013).

Another promising application for NILV is their use as templates for site-directed integration systems. A variety of systems are available that can direct integration to genomic “safe loci” or by altering the integration pattern to avoid transcriptional units with the hope of minimizing gene dysregulation. Integration can be

directed to sequence-specific motifs in less intragenic regions by combining NILVs with an integrase protein fused to a DNA-binding protein such as the E. coli LexA repressor (Goulaouic & Chow, 1996; Holmes-Son & Chow, 2000, 2002) or a synthetic polydactyl zinc finger protein E2C (Tan, Dong, Wilkinson, Barbas, & Chow, 2006). Another approach is combining recombinases or transposases transiently with NILV to facilitate integration at specific sites (Moldt et al., 2011; Moldt, Staunstrup, Jakobsen, Yanez-Munoz, & Mikkelsen, 2008; Vink et al., 2009). Third, NILVs can be designed to promote site specific homologous recombination (HR) (Okada et al., 2009). Taking this a step farther, others have combined NILV with a rare cutting nuclease for targeted recombination at specific sites by HR (Cornu & Cathomen, 2007). Still others have used NILV as templates for HR along with engineered zinc finger nucleases (ZFNs) or transcription activator-like effector nucleases (TALENs) (Joglekar et al., 2013; Lombardo et al., 2011; Lombardo et al., 2007; Osborn et al., 2013). Increasing clinical utility is expected as these systems are optimized to reduce off-target integrations and increase the efficiency of delivery.

NILV have potential utility in cytotoxic cancer therapies. One approach takes advantage of abnormal expression levels of miRNA which are found in many tumor types (Calin & Croce, 2006). In vitro and in vivo studies have shown some miRNAs to have antitumorigenic properties (Ventura & Jacks, 2009). Recent studies have shown the utility of non-integrating adeno-associated virus in suppressing tumor growth in lung (Esquela-Kerscher et al., 2008; Kumar et al., 2008) and liver cancers (Hsu et al., 2012; Kota et al., 2009). Non-integrating vectors offer the advantage of minimizing effects from transgene expression in normal cells. This targeting of miRNA provides the

backdrop for developing NILV for similar applications, whose lower levels of transgene expression may be advantageous as too much miRNA can inadvertently kill off-target cells. We have initiated the development of an NILV for these applications containing our novel modifications. We have inserted a microRNA into our novel vector combination in order to test its utility for cytotoxic cancer therapy of the liver. RT-Q-PCR analysis demonstrated that our vector was able to provide approximately a 50X fold increase in miR-122 expression in HEK 293 cells. Initial testing was initiated in various hepatic carcinoma cell lines including Hep-G2, Hep-3B, SNU-449 and Hepa-1c1c7.

While there are many applications for transient expression of NILV, there is also great promise in their utilization for persistent episomal expression in non-proliferating post-mitotic cells. In this regard, NILV have been found to successfully provide long-lasting in vivo expression in several organs. Injections of integrase deficient LV and att site mutant NILV into mouse muscle were found to provide levels of transgene expression similar to wild-type LV for up three months post-transduction (Apolonia et al., 2007). NILV have also been used to transduce the liver resulting in stable transgene expression for up to six months (Bayer et al., 2008) and could provide therapeutic levels of transgene expression (Cantore et al., 2012; Matrai et al., 2011). The retina transduced with NILVs has been shown to provide transgene expression for up to nine months in mice (Yanez-Munoz et al., 2006). NILV were used to successfully transduce the brain and spinal cord allowing for efficient transgene expression from 2 weeks up to 4 months post-transduction (Kantor et al., 2011; Peluffo et al., 2013; Philippe et al., 2006; Rahim et al., 2009; Yanez-Munoz et al., 2006). This list is by no

means conclusive and as new technologies are developed optimized NILV will continue to provide a safe, flexible platform for clinical applications.

REFERENCES

Aiken, C. (1997). Pseudotyping human immunodeficiency virus type 1 (HIV-1) by the glycoprotein of vesicular stomatitis virus targets HIV-1 entry to an endocytic pathway and suppresses both the requirement for Nef and the sensitivity to cyclosporin A. *J Virol, 71*(8), 5871-5877.

Aiuti, A., Biasco, L., Scaramuzza, S., Ferrua, F., Cicalese, M. P., Baricordi, C., . . . Naldini, L. (2013). Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science, 341*(6148), 1233151. doi: 10.1126/science.1233151

Apolonia, L., Waddington, S. N., Fernandes, C., Ward, N. J., Bouma, G., Blundell, M. P., . . . Philpott, N. J. (2007). Stable gene transfer to muscle using non-integrating lentiviral vectors. *Mol Ther, 15*(11), 1947-1954. doi: 10.1038/sj.mt.6300281

Banasik, M. B., & McCray, P. B., Jr. (2010). Integrase-defective lentiviral vectors: progress and applications. *Gene Ther, 17*(2), 150-157. doi: 10.1038/gt.2009.135

Barry, S. C., Harder, B., Brzezinski, M., Flint, L. Y., Seppen, J., & Osborne, W. R. (2001). Lentivirus vectors encoding both central polypurine tract and posttranscriptional regulatory element provide enhanced transduction and transgene expression. *Hum Gene Ther, 12*(9), 1103-1108. doi: 10.1089/104303401750214311

Bayer, M., Kantor, B., Cockrell, A., Ma, H., Zeithaml, B., Li, X., . . . Kafri, T. (2008). A large U3 deletion causes increased in vivo expression from a nonintegrating lentiviral vector. *Mol Ther, 16*(12), 1968-1976. doi: 10.1038/mt.2008.199

Berger, G., Goujon, C., Darlix, J. L., & Cimarelli, A. (2009). SIVMAC Vpx improves the transduction of dendritic cells with nonintegrative HIV-1-derived vectors. *Gene Ther, 16*(1), 159-163. doi: 10.1038/gt.2008.128

Beumer, K., Bhattacharyya, G., Bibikova, M., Trautman, J. K., & Carroll, D. (2006). Efficient gene targeting in Drosophila with zinc-finger nucleases. *Genetics, 172*(4), 2391-2403. doi: 10.1534/genetics.105.052829

Biffi, A., Montini, E., Lorioli, L., Cesani, M., Fumagalli, F., Plati, T., . . . Naldini, L. (2013). Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science, 341*(6148), 1233158. doi: 10.1126/science.1233158

Braun, C. J., Boztug, K., Paruzynski, A., Witzel, M., Schwarzer, A., Rothe, M., . . . Klein, C. (2014). Gene therapy for Wiskott-Aldrich syndrome--long-term efficacy and genotoxicity. *Sci Transl Med, 6*(227), 227ra233. doi: 10.1126/scitranslmed.3007280

Brown, H. E., Chen, H., & Engelman, A. (1999). Structure-based mutagenesis of the human immunodeficiency virus type 1 DNA attachment site: effects on integration and cDNA synthesis. *J Virol, 73*(11), 9011-9020.

Bukrinsky, M. I., Sharova, N., Dempsey, M. P., Stanwick, T. L., Bukrinskaya, A. G., Haggerty, S., & Stevenson, M. (1992). Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. *Proc Natl Acad Sci U S A, 89*(14), 6580-6584.

Bukrinsky, M. I., Sharova, N., McDonald, T. L., Pushkarskaya, T., Tarpley, W. G., & Stevenson, M. (1993). Association of integrase, matrix, and reverse transcriptase

antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc Natl Acad Sci U S A, 90*(13), 6125-6129.

Calin, G. A., & Croce, C. M. (2006). MicroRNA signatures in human cancers. *Nat Rev Cancer, 6*(11), 857-866. doi: 10.1038/nrc1997

Cantore, A., Nair, N., Della Valle, P., Di Matteo, M., Matrai, J., Sanvito, F., . . . Vandendriessche, T. (2012). Hyperfunctional coagulation factor IX improves the efficacy of gene therapy in hemophilic mice. *Blood, 120*(23), 4517-4520. doi: 10.1182/blood-2012-05-432591

Cartier, N., Hacein-Bey-Abina, S., Bartholomae, C. C., Veres, G., Schmidt, M., Kutschera, I., . . . Aubourg, P. (2009). Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science, 326*(5954), 818-823. doi: 10.1126/science.1171242

Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., . . . Leboulch, P. (2010). Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. *Nature, 467*(7313), 318-322. doi: 10.1038/nature09328

Chun, T. W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J. A., Taylor, H., . . . Siliciano, R. F. (1997). Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature, 387*(6629), 183-188. doi: 10.1038/387183a0

Clapham, P. R., & McKnight, A. (2002). Cell surface receptors, virus entry and tropism of primate lentiviruses. *J Gen Virol, 83*(Pt 8), 1809-1829.

Cornetta, K., Yao, J., Jasti, A., Koop, S., Douglas, M., Hsu, D., . . . Duffy, L. (2011). Replication-competent lentivirus analysis of clinical grade vector products. *Mol Ther, 19*(3), 557-566. doi: 10.1038/mt.2010.278

Cornu, T. I., & Cathomen, T. (2007). Targeted genome modifications using integrase-deficient lentiviral vectors. *Mol Ther, 15*(12), 2107-2113. doi: 10.1038/sj.mt.6300345

Coutant, F., Frenkiel, M. P., Despres, P., & Charneau, P. (2008). Protective antiviral immunity conferred by a nonintegrative lentiviral vector-based vaccine. *PLoS One, 3*(12), e3973. doi: 10.1371/journal.pone.0003973

Coutant, F., Sanchez David, R. Y., Felix, T., Boulay, A., Caleechurn, L., Souque, P., . . . Charneau, P. (2012). A nonintegrative lentiviral vector-based vaccine provides long-term sterile protection against malaria. *PLoS One, 7*(11), e48644. doi: 10.1371/journal.pone.0048644

Craigie, R., Fujiwara, T., & Bushman, F. (1990). The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. *Cell, 62*(4), 829-837.

Daenthanasanmak, A., Salguero, G., Borchers, S., Figueiredo, C., Jacobs, R., Sundarasetty, B. S., . . . Stripecke, R. (2012). Integrase-defective lentiviral vectors encoding cytokines induce differentiation of human dendritic cells and stimulate multivalent immune responses in vitro and in vivo. *Vaccine, 30*(34), 5118-5131. doi: 10.1016/j.vaccine.2012.05.063

De Ravin, S. S., Su, L., Theobald, N., Choi, U., Macpherson, J. L., Poidinger, M., . . . Wu, X. (2014). Enhancers are major targets for murine leukemia virus vector integration. *J Virol, 88*(8), 4504-4513. doi: 10.1128/JVI.00011-14

De Rijck, J., Vandekerckhove, L., Christ, F., & Debyser, Z. (2007). Lentiviral nuclear import: a complex interplay between virus and host. *Bioessays, 29*(5), 441-451. doi: 10.1002/bies.20561

Deng, Y., Guan, J., Wen, B., Zhu, N., Chen, H., Song, J., . . . Tan, W. (2013). Induction of broadly neutralising HCV antibodies in mice by integration-deficient lentiviral vector-based pseudotyped particles. *PLoS One, 8*(4), e62684. doi: 10.1371/journal.pone.0062684

DiGiusto, D. L., Krishnan, A., Li, L., Li, H., Li, S., Rao, A., . . . Zaia, J. A. (2010). RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma. *Sci Transl Med, 2*(36), 36ra43. doi: 10.1126/scitranslmed.3000931

Dina, D., & Benz, E. W., Jr. (1980). Structure of murine sarcoma virus DNA replicative intermediates synthesized in vitro. *J Virol, 33*(1), 377-389.

Dull, T., Zufferey, R., Kelly, M., Mandel, R. J., Nguyen, M., Trono, D., & Naldini, L. (1998). A third-generation lentivirus vector with a conditional packaging system. *J Virol, 72*(11), 8463-8471.

Dupuy, F. P., Mouly, E., Mesel-Lemoine, M., Morel, C., Abriol, J., Cherai, M., . . . Lemoine, F. M. (2005). Lentiviral transduction of human hematopoietic cells by HIV-1- and

SIV-based vectors containing a bicistronic cassette driven by various internal promoters. *J Gene Med, 7*(9), 1158-1171. doi: 10.1002/jgm.769

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics, 26*(19), 2460-2461. doi: 10.1093/bioinformatics/btq461

Engelman, A. (1999). In vivo analysis of retroviral integrase structure and function. *Adv Virus Res, 52*, 411-426.

Esposito, D., & Craigie, R. (1998). Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein-DNA interaction. *EMBO J, 17*(19), 5832-5843. doi: 10.1093/emboj/17.19.5832

Esquela-Kerscher, A., Trang, P., Wiggins, J. F., Patrawala, L., Cheng, A., Ford, L., . . . Slack, F. J. (2008). The let-7 microRNA reduces tumor growth in mouse models of lung cancer. *Cell Cycle, 7*(6), 759-764.

Farnet, C. M., & Haseltine, W. A. (1991a). Circularization of human immunodeficiency virus type 1 DNA in vitro. *J Virol, 65*(12), 6942-6952.

Farnet, C. M., & Haseltine, W. A. (1991b). Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. *J Virol, 65*(4), 1910-1915.

Fassati, A. (2006). HIV infection of non-dividing cells: a divisive problem. *Retrovirology, 3*, 74. doi: 10.1186/1742-4690-3-74

Fassati, A., & Goff, S. P. (2001). Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1. *J Virol, 75*(8), 3626-3635. doi: 10.1128/JVI.75.8.3626-3635.2001

Fischer, U., Huber, J., Boelens, W. C., Mattaj, I. W., & Luhrmann, R. (1995). The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell, 82*(3), 475-483.

Folger, K. R., Wong, E. A., Wahl, G., & Capecchi, M. R. (1982). Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. *Mol Cell Biol, 2*(11), 1372-1387.

Freed, E. O. (1998). HIV-1 gag proteins: diverse functions in the virus life cycle. *Virology, 251*(1), 1-15. doi: 10.1006/viro.1998.9398

Gabriel, R., Eckenberg, R., Paruzynski, A., Bartholomae, C. C., Nowrouzi, A., Arens, A., . . . Schmidt, M. (2009). Comprehensive genomic access to vector integration in clinical gene therapy. *Nat Med, 15*(12), 1431-1436. doi: 10.1038/nm.2057

Gallay, P., Hope, T., Chin, D., & Trono, D. (1997). HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. *Proc Natl Acad Sci U S A, 94*(18), 9825-9830.

Gasmi, M., Glynn, J., Jin, M. J., Jolly, D. J., Yee, J. K., & Chen, S. T. (1999). Requirements for efficient production and transduction of human immunodeficiency virus type 1-based vectors. *J Virol, 73*(3), 1828-1834.

Gaur, M., & Leavitt, A. D. (1998). Mutations in the human immunodeficiency virus type 1 integrase D,D(35)E motif do not eliminate provirus formation. *J Virol, 72*(6), 4678-4685.

Gerton, J. L., Ohgi, S., Olsen, M., DeRisi, J., & Brown, P. O. (1998). Effects of mutations in residues near the active site of human immunodeficiency virus type 1 integrase on specific enzyme-substrate interactions. *J Virol, 72*(6), 5046-5055.

Gianni, A. M., Smotkin, D., & Weinberg, R. A. (1975). Murine leukemia virus: detection of unintegrated double-stranded DNA forms of the provirus. *Proc Natl Acad Sci U S A, 72*(2), 447-451.

Gilboa, E., Mitra, S. W., Goff, S., & Baltimore, D. (1979). A detailed model of reverse transcription and tests of crucial aspects. *Cell, 18*(1), 93-100.

Goulaouic, H., & Chow, S. A. (1996). Directed integration of viral DNA mediated by fusion proteins consisting of human immunodeficiency virus type 1 integrase and Escherichia coli LexA protein. *J Virol, 70*(1), 37-46.

Grasso, F., Negri, D. R., Mochi, S., Rossi, A., Cesolini, A., Giovannelli, A., . . . Cara, A. (2012). Successful therapeutic vaccination with integrase defective lentiviral vector expressing nononcogenic human papillomavirus E7 protein. *Int J Cancer*. doi: 10.1002/ijc.27676

Hacein-Bey-Abina, S., Garrigue, A., Wang, G. P., Soulier, J., Lim, A., Morillon, E., . . . Cavazzana-Calvo, M. (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J Clin Invest, 118*(9), 3132-3142. doi: 10.1172/JCI35700

Hawkins, T. B., Dantzer, J., Peters, B., Dinauer, M., Mockaitis, K., Mooney, S., & Cornetta, K. (2011). Identifying viral integration sites using SeqMap 2.0. *Bioinformatics, 27*(5), 720-722. doi: 10.1093/bioinformatics/btq722

Higashikawa, F., & Chang, L. (2001). Kinetic analyses of stability of simple and complex retroviral vectors. *Virology, 280*(1), 124-131. doi: 10.1006/viro.2000.0743

Hindmarsh, P., & Leis, J. (1999). Retroviral DNA integration. *Microbiol Mol Biol Rev, 63*(4), 836-843, table of contents.

Holmes-Son, M. L., & Chow, S. A. (2000). Integrase-lexA fusion proteins incorporated into human immunodeficiency virus type 1 that contains a catalytically inactive integrase gene are functional to mediate integration. *J Virol, 74*(24), 11548-11556.

Holmes-Son, M. L., & Chow, S. A. (2002). Correct integration mediated by integrase-LexA fusion proteins incorporated into HIV-1. *Mol Ther, 5*(4), 360-370. doi: 10.1006/mthe.2002.0559

Hoover, T., Mikovits, J., Court, D., Liu, Y. L., Kung, H. F., & Raziuddin. (1996). A nuclear matrix-specific factor that binds a specific segment of the negative regulatory element (NRE) of HIV-1 LTR and inhibits NF-kappa(B) activity. *Nucleic Acids Res, 24*(10), 1895-1900.

Howe, S. J., Mansour, M. R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempski, H., . . . Thrasher, A. J. (2008). Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J Clin Invest, 118*(9), 3143-3150. doi: 10.1172/JCI35798

Hsu, S. H., Wang, B., Kota, J., Yu, J., Costinean, S., Kutay, H., . . . Ghoshal, K. (2012). Essential metabolic, anti-inflammatory, and anti-tumorigenic functions of miR-122 in liver. *J Clin Invest, 122*(8), 2871-2883. doi: 10.1172/JCI63539

Hu, B., Dai, B., & Wang, P. (2010). Vaccines delivered by integration-deficient lentiviral vectors targeting dendritic cells induces strong antigen-specific immunity. *Vaccine, 28*(41), 6675-6683. doi: 10.1016/j.vaccine.2010.08.012

Jacque, J. M., & Stevenson, M. (2006). The inner-nuclear-envelope protein emerin regulates HIV-1 infectivity. *Nature, 441*(7093), 641-645. doi: 10.1038/nature04682

Jeanson, L., Subra, F., Vaganay, S., Hervy, M., Marangoni, E., Bourhis, J., & Mouscadet, J. F. (2002). Effect of Ku80 depletion on the preintegrative steps of HIV-1 replication in human cells. *Virology, 300*(1), 100-108.

Joglekar, A. V., Hollis, R. P., Kuftinec, G., Senadheera, S., Chan, R., & Kohn, D. B. (2013). Integrase-defective lentiviral vectors as a delivery platform for targeted modification of adenosine deaminase locus. *Mol Ther, 21*(9), 1705-1717. doi: 10.1038/mt.2013.106

Johnson, A. J., Goger, P. R., & Tillett, W. S. (1954). The intravenous injection of bovine crystalline pancreatic desoxyribonuclease into patients. *J Clin Invest, 33*(12), 1670-1686. doi: 10.1172/JCI103048

Jordan, M., & Wurm, F. (2004). Transfection of adherent and suspended cells by calcium phosphate. *Methods, 33*(2), 136-143. doi: 10.1016/j.ymeth.2003.11.011

Junghans, R. P., Boone, L. R., & Skalka, A. M. (1982). Products of reverse transcription in avian retrovirus analyzed by electron microscopy. *J Virol, 43*(2), 544-554.

Kalos, M., Levine, B. L., Porter, D. L., Katz, S., Grupp, S. A., Bagg, A., & June, C. H. (2011). T cells with chimeric antigen receptors have potent antitumor effects and can

establish memory in patients with advanced leukemia. *Sci Transl Med, 3*(95), 95ra73. doi: 10.1126/scitranslmed.3002842

Kantor, B., Bayer, M., Ma, H., Samulski, J., Li, C., McCown, T., & Kafri, T. (2011). Notable reduction in illegitimate integration mediated by a PPT-deleted, nonintegrating lentiviral vector. *Mol Ther, 19*(3), 547-556. doi: 10.1038/mt.2010.277

Karwacz, K., Mukherjee, S., Apolonia, L., Blundell, M. P., Bouma, G., Escors, D., . . . Thrasher, A. J. (2009). Nonintegrating lentivector vaccines stimulate prolonged T-cell and antibody responses and are effective in tumor therapy. *J Virol, 83*(7), 3094-3103. doi: 10.1128/JVI.02519-08

Katz, R. A., Merkel, G., Kulkosky, J., Leis, J., & Skalka, A. M. (1990). The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. *Cell, 63*(1), 87-95.

Kilzer, J. M., Stracker, T., Beitzel, B., Meek, K., Weitzman, M., & Bushman, F. D. (2003). Roles of host cell factors in circularization of retroviral dna. *Virology, 314*(1), 460-467.

Klarmann, G. J., Yu, H., Chen, X., Dougherty, J. P., & Preston, B. D. (1997). Discontinuous plus-strand DNA synthesis in human immunodeficiency virus type 1-infected cells and in a partially reconstituted cell-free system. *J Virol, 71*(12), 9259-9269.

Klaver, B., & Berkhout, B. (1994). Comparison of 5' and 3' long terminal repeat promoter function in human immunodeficiency virus. *J Virol, 68*(6), 3830-3840.

Kota, J., Chivukula, R. R., O'Donnell, K. A., Wentzel, E. A., Montgomery, C. L., Hwang, H. W., . . . Mendell, J. T. (2009). Therapeutic microRNA delivery suppresses

tumorigenesis in a murine liver cancer model. *Cell, 137*(6), 1005-1017. doi: 10.1016/j.cell.2009.04.021

Koyama, T., Sun, B., Tokunaga, K., Tatsumi, M., & Ishizaka, Y. (2013). DNA damage enhances integration of HIV-1 into macrophages by overcoming integrase inhibition. *Retrovirology, 10*, 21. doi: 10.1186/1742-4690-10-21

Kulkosky, J., Jones, K. S., Katz, R. A., Mack, J. P., & Skalka, A. M. (1992). Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. *Mol Cell Biol, 12*(5), 2331-2338.

Kulkosky, J., & Skalka, A. M. (1994). Molecular mechanism of retroviral DNA integration. *Pharmacol Ther, 61*(1-2), 185-203.

Kumar, M. S., Erkeland, S. J., Pester, R. E., Chen, C. Y., Ebert, M. S., Sharp, P. A., & Jacks, T. (2008). Suppression of non-small cell lung tumor development by the let-7 microRNA family. *Proc Natl Acad Sci U S A, 105*(10), 3903-3908. doi: 10.1073/pnas.0712321105

Lachmann, P. J. (1967). Allergic reactions, connective tissue, and disease. *Sci Basis Med Annu Rev*, 36-58.

Leath, A., & Cornetta, K. (2012). Developing novel lentiviral vectors into clinical products. . *Methods in Enzymology, 507*, 89-108.

Leavitt, A. D., Robles, G., Alesandro, N., & Varmus, H. E. (1996). Human immunodeficiency virus type 1 integrase mutants retain in vitro integrase activity yet fail to integrate viral DNA efficiently during infection. *J Virol, 70*(2), 721-728.

Leavitt, A. D., Shiue, L., & Varmus, H. E. (1993). Site-directed mutagenesis of HIV-1 integrase demonstrates differential effects on integrase functions in vitro. *J Biol Chem, 268*(3), 2113-2119.

Levine, B. L., Humeau, L. M., Boyer, J., MacGregor, R. R., Rebello, T., Lu, X., . . . June, C. H. (2006). Gene transfer in humans using a conditionally replicating lentiviral vector. *Proc Natl Acad Sci U S A, 103*(46), 17372-17377. doi: 10.1073/pnas.0608138103

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics, 25*(14), 1754-1760. doi: 10.1093/bioinformatics/btp324

Li, L., Olvera, J. M., Yoder, K. E., Mitchell, R. S., Butler, S. L., Lieber, M., . . . Bushman, F. D. (2001). Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection. *EMBO J, 20*(12), 3272-3281. doi: 10.1093/emboj/20.12.3272

Lombardo, A., Cesana, D., Genovese, P., Di Stefano, B., Provasi, E., Colombo, D. F., . . . Naldini, L. (2011). Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat Methods, 8*(10), 861-869. doi: 10.1038/nmeth.1674

Lombardo, A., Genovese, P., Beausejour, C. M., Colleoni, S., Lee, Y. L., Kim, K. A., . . . Naldini, L. (2007). Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol, 25*(11), 1298-1306. doi: 10.1038/nbt1353

Loyter, A., Scangos, G. A., & Ruddle, F. H. (1982). Mechanisms of DNA uptake by mammalian cells: fate of exogenously added DNA monitored by the use of fluorescent dyes. *Proc Natl Acad Sci U S A, 79*(2), 422-426.

Lu, X., Humeau, L., Slepushkin, V., Binder, G., Yu, Q., Slepushkina, T., . . . Dropulic, B. (2004). Safe two-plasmid production for the first clinical lentivirus vector that achieves >99% transduction in primary cells using a one-step protocol. *J Gene Med, 6*(9), 963-973. doi: 10.1002/jgm.593

Lu, Y., Stenzel, M., Sodroski, J. G., & Haseltine, W. A. (1989). Effects of long terminal repeat mutations on human immunodeficiency virus type 1 replication. *J Virol, 63*(9), 4115-4119.

Mali, P., Esvelt, K. M., & Church, G. M. (2013). Cas9 as a versatile tool for engineering biology. *Nat Methods, 10*(10), 957-963. doi: 10.1038/nmeth.2649

Mali, P., Ye, Z., Hommond, H. H., Yu, X., Lin, J., Chen, G., . . . Cheng, L. (2008). Improved efficiency and pace of generating induced pluripotent stem cells from human adult and fetal fibroblasts. *Stem Cells, 26*(8), 1998-2005. doi: 10.1634/stemcells.2008-0346

Malim, M. H., & Emerman, M. (2008). HIV-1 accessory proteins--ensuring viral survival in a hostile environment. *Cell Host Microbe, 3*(6), 388-398. doi: 10.1016/j.chom.2008.04.008

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal, 17*(1), pp. 10-12.

Masuda, T., Kuroda, M. J., & Harada, S. (1998). Specific and independent recognition of U3 and U5 att sites by human immunodeficiency virus type 1 integrase in vivo. *J Virol, 72*(10), 8396-8402.

Masuda, T., Planelles, V., Krogstad, P., & Chen, I. S. (1995). Genetic analysis of human immunodeficiency virus type 1 integrase and the U3 att site: unusual phenotype of mutants in the zinc finger-like domain. *J Virol, 69*(11), 6687-6696.

Matrai, J., Cantore, A., Bartholomae, C. C., Annoni, A., Wang, W., Acosta-Sanchez, A., . . . Naldini, L. (2011). Hepatocyte-targeted expression by integrase-defective lentiviral vectors induces antigen-specific tolerance in mice with low genotoxic risk. *Hepatology, 53*(5), 1696-1707. doi: 10.1002/hep.24230

Matrai, J., Chuah, M. K., & VandenDriessche, T. (2010). Recent advances in lentiviral vector development and applications. *Mol Ther, 18*(3), 477-490. doi: 10.1038/mt.2009.319

McBride, M. S., & Panganiban, A. T. (1996). The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures. *J Virol, 70*(5), 2963-2973.

Merten, O., Charrier, S., Laroudie, N., Fauchille, S., Dugue, C., Jenny, C., . . . Galy, A. (2011). Large scale manufacture and characterisation of a lentiviral vector produced for clinical ex vivo gene therapy application. *Hum Gene Ther, 22*, 343-356.

Michelini, Z., Negri, D. R., Baroncelli, S., Spada, M., Leone, P., Bona, R., . . . Cara, A. (2009). Development and use of SIV-based Integrase defective lentiviral vector

for immunization. *Vaccine, 27*(34), 4622-4629. doi: 10.1016/j.vaccine.2009.05.070

Miller, D. G., Petek, L. M., & Russell, D. W. (2004). Adeno-associated virus vectors integrate at chromosome breakage sites. *Nat Genet, 36*(7), 767-773. doi: 10.1038/ng1380

Miller, J. C., Holmes, M. C., Wang, J., Guschin, D. Y., Lee, Y. L., Rupniewski, I., . . . Rebar, E. J. (2007). An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol, 25*(7), 778-785. doi: 10.1038/nbt1319

Miller, M. D., Farnet, C. M., & Bushman, F. D. (1997). Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *J Virol, 71*(7), 5382-5390.

Miyoshi, H., Blomer, U., Takahashi, M., Gage, F. H., & Verma, I. M. (1998). Development of a self-inactivating lentivirus vector. *J Virol, 72*(10), 8150-8157.

Mizuuchi, K. (1992). Polynucleotidyl transfer reactions in transpositional DNA recombination. *J Biol Chem, 267*(30), 21273-21276.

Modlich, U., Navarro, S., Zychlinski, D., Maetzig, T., Knoess, S., Brugman, M. H., . . . Baum, C. (2009). Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. *Mol Ther, 17*(11), 1919-1928. doi: 10.1038/mt.2009.179

Moldt, B., Miskey, C., Staunstrup, N. H., Gogol-Doring, A., Bak, R. O., Sharma, N., . . . Mikkelsen, J. G. (2011). Comparative genomic integration profiling of Sleeping Beauty transposons mobilized with high efficacy from integrase-defective

lentiviral vectors in primary human cells. *Mol Ther, 19*(8), 1499-1510. doi: 10.1038/mt.2011.47

Moldt, B., Staunstrup, N. H., Jakobsen, M., Yanez-Munoz, R. J., & Mikkelsen, J. G. (2008). Genomic insertion of lentiviral DNA circles directed by the yeast Flp recombinase. *BMC Biotechnol, 8*, 60. doi: 10.1186/1472-6750-8-60

Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Ponzoni, M., Bartholomae, C., . . . Naldini, L. (2006). Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nat Biotechnol, 24*(6), 687-696. doi: 10.1038/nbt1216

Mueller, P. R., & Wold, B. (1989). In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. *Science, 246*(4931), 780-786.

Nakajima, N., Lu, R., & Engelman, A. (2001). Human immunodeficiency virus type 1 replication in the absence of integrase-mediated dna recombination: definition of permissive and nonpermissive T-cell lines. *J Virol, 75*(17), 7944-7955.

Naldini, L., Blomer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F. H., . . . Trono, D. (1996). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science, 272*(5259), 263-267.

Negri, D. R., Michelini, Z., Baroncelli, S., Spada, M., Vendetti, S., Buffa, V., . . . Cara, A. (2007). Successful immunization with a single injection of non-integrating lentiviral vector. *Mol Ther, 15*(9), 1716-1723. doi: 10.1038/sj.mt.6300241

Negri, D. R., Rossi, A., Blasi, M., Michelini, Z., Leone, P., Chiantore, M. V., . . . Cara, A. (2012). Simian immunodeficiency virus-Vpx for improving integrase defective

lentiviral vector-based vaccines. *Retrovirology, 9*, 69. doi: 10.1186/1742-4690-9-69

Nightingale, S. J., Hollis, R. P., Pepper, K. A., Petersen, D., Yu, X. J., Yang, C., . . . Kohn, D. B. (2006). Transient gene expression by nonintegrating lentiviral vectors. *Mol Ther, 13*(6), 1121-1132. doi: 10.1016/j.ymthe.2006.01.008

Okada, Y., Ueshin, Y., Hasuwa, H., Takumi, K., Okabe, M., & Ikawa, M. (2009). Targeted gene modification in mouse ES cells using integrase-defective lentiviral vectors. *Genesis, 47*(4), 217-223. doi: 10.1002/dvg.20469

Okita, K., Ichisaka, T., & Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. *Nature, 448*(7151), 313-317. doi: 10.1038/nature05934

Osborn, M. J., Starker, C. G., McElroy, A. N., Webber, B. R., Riddle, M. J., Xia, L., . . . Tolar, J. (2013). TALEN-based gene correction for epidermolysis bullosa. *Mol Ther, 21*(6), 1151-1159. doi: 10.1038/mt.2013.56

Ott, M. G., Schmidt, M., Schwarzwaelder, K., Stein, S., Siler, U., Koehl, U., . . . Grez, M. (2006). Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat Med, 12*(4), 401-409. doi: 10.1038/nm1393

Pan, C. Q., Sinicropi, D. V., & Lazarus, R. A. (2001). Engineered properties and assays for human DNase I mutants. *Methods Mol Biol, 160*, 309-321. doi: 10.1385/1-59259-233-3:309

Papapetrou, E. P., & Sadelain, M. (2011). Generation of transgene-free human induced pluripotent stem cells with an excisable single polycistronic vector. *Nat Protoc, 6*(9), 1251-1273. doi: 10.1038/nprot.2011.374

Papapetrou, E. P., Tomishima, M. J., Chambers, S. M., Mica, Y., Reed, E., Menon, J., . . . Sadelain, M. (2009). Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. *Proc Natl Acad Sci U S A, 106*(31), 12759-12764. doi: 10.1073/pnas.0904825106

Pelascini, L. P., Janssen, J. M., & Goncalves, M. A. (2013). Histone deacetylase inhibition activates transgene expression from integration-defective lentiviral vectors in dividing and non-dividing cells. *Hum Gene Ther, 24*(1), 78-96. doi: 10.1089/hum.2012.069

Peluffo, H., Foster, E., Ahmed, S. G., Lago, N., Hutson, T. H., Moon, L., . . . Yanez-Munoz, R. J. (2013). Efficient gene expression from integration-deficient lentiviral vectors in the spinal cord. *Gene Ther, 20*(6), 645-657. doi: 10.1038/gt.2012.78

Philippe, S., Sarkis, C., Barkats, M., Mammeri, H., Ladroue, C., Petit, C., . . . Serguera, C. (2006). Lentiviral vectors with a defective integrase allow efficient and sustained transgene expression in vitro and in vivo. *Proc Natl Acad Sci U S A, 103*(47), 17684-17689. doi: 10.1073/pnas.0606197103

Porter, D. L., Levine, B. L., Kalos, M., Bagg, A., & June, C. H. (2011). Chimeric Antigen Receptor-Modified T Cells in Chronic Lymphoid Leukemia. *New England Journal of Medicine, 365*, 725-733.

Rahim, A. A., Wong, A. M., Howe, S. J., Buckley, S. M., Acosta-Saltos, A. D., Elston, K. E., . . . Raivich, G. (2009). Efficient gene delivery to the adult and fetal CNS using pseudotyped non-integrating lentiviral vectors. *Gene Ther, 16*(4), 509-520. doi: 10.1038/gt.2008.186

Ramezani, A., Hawley, T. S., & Hawley, R. G. (2008). Combinatorial incorporation of enhancer-blocking components of the chicken beta-globin 5'HS4 and human T-cell receptor alpha/delta BEAD-1 insulators in self-inactivating retroviral vectors reduces their genotoxic potential. *Stem Cells, 26*(12), 3257-3266. doi: 10.1634/stemcells.2008-0258

Raskin, P. (1968). Bronchospasm after inhalation of pancreatic dornase. *Am Rev Respir Dis, 98*(4), 697-698.

RCore, T. (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria: ISBN 3-900051-07-0, URL http://www. R-project. org.

Ringold, G. M., Yamamoto, K. R., Shank, P. R., & Varmus, H. E. (1977). Mouse mammary tumor virus DNA in infected rat cells: characterization of unintegrated forms. *Cell, 10*(1), 19-26.

Roche. (2014). Pulmoyzme (dornase alfa) Product Information. http://www.roche-australia.com/content/dam/internet/corporate/roche/en_AU/files/miscellaneous/pulmozyme-pi.pdf

Rosen, C. A., Sodroski, J. G., & Haseltine, W. A. (1985). The location of cis-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. *Cell, 41*(3), 813-823.

Saenz, D. T., Loewen, N., Peretz, M., Whitwam, T., Barraza, R., Howell, K. G., . . . Poeschla, E. M. (2004). Unintegrated lentivirus DNA persistence and accessibility to expression in nondividing cells: analysis with class I integrase mutants. *J Virol, 78*(6), 2906-2920.

Sastry, L., Johnson, T., Hobson, M. J., Smucker, B., & Cornetta, K. (2002). Titering lentiviral vectors: comparison of DNA, RNA and marker expression methods. *Gene Ther, 9*(17), 1155-1162. doi: 10.1038/sj.gt.3301731

Sastry, L., Xu, Y., Cooper, R., Pollok, K., & Cornetta, K. (2004). Evaluation of plasmid DNA removal from lentiviral vectors by benzonase treatment. *Hum Gene Ther, 15*(2), 221-226. doi: 10.1089/104303404772680029

Sastry, L., Xu, Y., Johnson, T., Desai, K., Rissing, D., Marsh, J., & Cornetta, K. (2003). Certification assays for HIV-1-based vectors: frequent passage of gag sequences without evidence of replication-competent viruses. *Mol Ther, 8*(5), 830-839.

Schmidt, M., Schwarzwaelder, K., Bartholomae, C., Zaoui, K., Ball, C., Pilz, I., . . . von Kalle, C. (2007). High-resolution insertion-site analysis by linear amplification-mediated PCR (LAM-PCR). *Nat Methods, 4*(12), 1051-1057. doi: 10.1038/nmeth1103

Schroder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., & Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell, 110*(4), 521-529.

Shak, S., Capon, D. J., Hellmiss, R., Marsters, S. A., & Baker, C. L. (1990). Recombinant human DNase I reduces the viscosity of cystic fibrosis sputum. *Proc Natl Acad Sci U S A, 87*(23), 9188-9192.

Shank, P. R., Hughes, S. H., Kung, H. J., Majors, J. E., Quintrell, N., Guntaka, R. V., . . . Varmus, H. E. (1978). Mapping unintegrated avian sarcoma virus DNA: termini of linear DNA bear 300 nucleotides present once or twice in two species of circular DNA. *Cell, 15*(4), 1383-1395.

Shibagaki, Y., & Chow, S. A. (1997). Central core domain of retroviral integrase is responsible for target site selection. *J Biol Chem, 272*(13), 8361-8369.

Shire, S. J. (1996). Stability characterization and formulation development of recombinant human deoxyribonuclease I [Pulmozyme, (dornase alpha)]. *Pharm Biotechnol, 9*, 393-426.

Shoemaker, C., Goff, S., Gilboa, E., Paskind, M., Mitra, S. W., & Baltimore, D. (1980). Structure of a cloned circular Moloney murine leukemia virus DNA molecule containing an inverted segment: implications for retrovirus integration. *Proc Natl Acad Sci U S A, 77*(7), 3932-3936.

Slepushkin, V., Chang, N., Cohen, R., Gan, Y., Jiang, B., Deausen, E., . . . Dropulic, B. (2003). Large-scale purification of a lentiviral vector by size exclusion

chromatography or mustang Q ion exchange capsule. *Bioprocessing Journal, Sept./Oct.*, 89-95.

Smith, D. R. (1992). Ligation-mediated PCR of restriction fragments from large DNA molecules. *PCR Methods Appl, 2*(1), 21-27.

Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G. W., Cook, E. G., . . . Jaenisch, R. (2009). Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. *Cell, 136*(5), 964-977. doi: 10.1016/j.cell.2009.02.013

Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G., & Hochedlinger, K. (2008). Induced pluripotent stem cells generated without viral integration. *Science, 322*(5903), 945-949. doi: 10.1126/science.1162494

Stevenson, M. (1996). Portals of entry: uncovering HIV nuclear transport pathways. *Trends Cell Biol, 6*(1), 9-15.

Suwanmanee, T., Hu, G., Gui, T., Bartholomae, C. C., Kutschera, I., von Kalle, C., . . . Kafri, T. (2013). Integration-deficient Lentiviral Vectors Expressing Codon-optimized R338L Human FIX Restore Normal Hemostasis in Hemophilia B Mice. *Mol Ther*. doi: 10.1038/mt.2013.188

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell, 131*(5), 861-872. doi: 10.1016/j.cell.2007.11.019

Tan, W., Dong, Z., Wilkinson, T. A., Barbas, C. F., 3rd, & Chow, S. A. (2006). Human immunodeficiency virus type 1 incorporated with fusion proteins consisting of

integrase and the designed polydactyl zinc finger protein E2C can bias integration of viral DNA into a predetermined chromosomal region in human cells. *J Virol, 80*(4), 1939-1948. doi: 10.1128/JVI.80.4.1939-1948.2006

Tareen, S. U., Kelley-Clarke, B., Nicolai, C. J., Cassiano, L. A., Nelson, L. T., Slough, M. M., . . . Robbins, S. H. (2014). Design of a novel integration-deficient lentivector technology that incorporates genetic and posttranslational elements to target human dendritic cells. *Mol Ther, 22*(3), 575-587. doi: 10.1038/mt.2013.278

Transfiguracion, J., Jaalouk, D. E., Ghani, K., Galipeau, J., & Kamen, A. (2003). Size-exclusion chromatography purification of high-titer vesicular stomatitis virus G glycoprotein-pseudotyped retrovectors for cell and gene therapy applications. *Hum Gene Ther, 14*(12), 1139-1153. doi: 10.1089/104303403322167984

Ulmer, J. S., Herzka, A., Toy, K. J., Baker, D. L., Dodge, A. H., Sinicropi, D., . . . Lazarus, R. A. (1996). Engineering actin-resistant human DNase I for treatment of cystic fibrosis. *Proc Natl Acad Sci U S A, 93*(16), 8225-8229.

Van Maele, B., De Rijck, J., De Clercq, E., & Debyser, Z. (2003). Impact of the central polypurine tract on the kinetics of human immunodeficiency virus type 1 vector transduction. *J Virol, 77*(8), 4685-4694.

Vargas, J., Jr., Gusella, G. L., Najfeld, V., Klotman, M. E., & Cara, A. (2004). Novel integrase-defective lentiviral episomal vectors for gene transfer. *Hum Gene Ther, 15*(4), 361-372. doi: 10.1089/104303404322959515

Vargas, J., Jr., Klotman, M. E., & Cara, A. (2008). Conditionally replicating lentiviral-hybrid episomal vectors for suicide gene therapy. *Antiviral Res, 80*(3), 288-294. doi: 10.1016/j.antiviral.2008.06.015

Ventura, A., & Jacks, T. (2009). MicroRNAs and cancer: short RNAs go a long way. *Cell, 136*(4), 586-591. doi: 10.1016/j.cell.2009.02.005

Vink, C. A., Gaspar, H. B., Gabriel, R., Schmidt, M., McIvor, R. S., Thrasher, A. J., & Qasim, W. (2009). Sleeping beauty transposition from nonintegrating lentivirus. *Mol Ther, 17*(7), 1197-1204. doi: 10.1038/mt.2009.94

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., . . . Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature, 448*(7151), 318-324. doi: 10.1038/nature05944

Wiskerchen, M., & Muesing, M. A. (1995). Human immunodeficiency virus type 1 integrase: effects of mutations on viral ability to integrate, direct viral gene expression from unintegrated viral DNA templates, and sustain viral propagation in primary cells. *J Virol, 69*(1), 376-386.

Wyatt, R., & Sodroski, J. (1998). The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science, 280*(5371), 1884-1888.

Yanez-Munoz, R. J., Balaggan, K. S., MacNeil, A., Howe, S. J., Schmidt, M., Smith, A. J., . . . Thrasher, A. J. (2006). Effective gene therapy with nonintegrating lentiviral vectors. *Nat Med, 12*(3), 348-353. doi: 10.1038/nm1365

Yang, G., Si-Tayeb, K., Corbineau, S., Vernet, R., Gayon, R., Dianat, N., . . . Weber, A. (2013). Integration-deficient lentivectors: an effective strategy to purify and

differentiate human embryonic stem cell-derived hepatic progenitors. *BMC Biol, 11*, 86. doi: 10.1186/1741-7007-11-86

Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, II, & Thomson, J. A. (2009). Human induced pluripotent stem cells free of vector and transgene sequences. *Science, 324*(5928), 797-801. doi: 10.1126/science.1172482

Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics, 30*(5), 614-620. doi: 10.1093/bioinformatics/btt593

Zhou, H., Rainey, G. J., Wong, S. K., & Coffin, J. M. (2001). Substrate sequence selection by retroviral integrase. *J Virol, 75*(3), 1359-1370. doi: 10.1128/JVI.75.3.1359-1370.2001

Zhou, S., Tang, C., Rapp, S., Bonner, M. A., De Ravin, S. S., Malech, H. L., & Sorrentino, B. P. (2014). *A Novel Non-Restrictive, Semi-Quantitative Method for Vector Insertion Site Analysis Based on Random Shearing of Genomic DNA.* Paper presented at the Molecular Therapy.

Zhu, K., Dobard, C., & Chow, S. A. (2004). Requirement for integrase during reverse transcription of human immunodeficiency virus type 1 and the effect of cysteine mutations of integrase on its interactions with reverse transcriptase. *J Virol, 78*(10), 5045-5055.

Zufferey, R. (2002). Production of lentiviral vectors. *Curr Top Microbiol Immunol, 261*, 107-121.

Zufferey, R., Donello, J. E., Trono, D., & Hope, T. J. (1999). Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *J Virol, 73*(4), 2886-2892.

Zufferey, R., Dull, T., Mandel, R. J., Bukovsky, A., Quiroz, D., Naldini, L., & Trono, D. (1998). Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J Virol, 72*(12), 9873-9880.

CURRICULUM VITAE

Aaron Marcus Shaw

Education

2009 – 2015 **Indiana University-Purdue University Indianapolis**, Indianapolis, IN
Ph.D. Medical and Molecular Genetics
Minor in Life Sciences
Dissertation: Advancing the Safety of Lentiviral Vector Mediated Gene Therapy
Advisor: Kenneth G. Cornetta, M.D.

2005 – 2009 **Indiana University Southeast**, New Albany, IN
B.S. Biology with Highest Distinction
Minor in Chemistry

2004 – 2005 **San Diego City College**, San Diego, CA
General Studies; Transferred

2002 – 2004 **Southwestern College**, Chula Vista, CA
General Studies; Transferred

1995 – 1996 **Jefferson Community College**, Louisville, KY
General Studies; Transferred

Professional Experience

2009 – 2015 **Graduate Research Assistant**. Indiana University School of Medicine, Department of Medical and Molecular Genetics, Indianapolis, IN.
Primary Investigator: Kenneth G. Cornetta, M.D.
1. Performed research to improve the safety of gene therapy.
2. Evaluated methods to improve lentiviral vector supernatant purity.
3. Designed and tested novel vectors for gene therapy applications.
4. Optimized protocols to improve evaluation of vector insertion sites.

2008 – 2009 **Undergraduate Research Fellow**. Indiana University Southeast, School of Natural Science, Biology Department, New Albany, IN.
Primary Investigator: Gretchen Kirchner, Ph.D.
1. Performed microbial genetics research.
2. Evaluated a novel bacteria mutant for applications in bioremediation.

2007 – 2009 **Laboratory Assistant**. Indiana University Southeast, School of Natural Science, Departments of Chemistry and Biology, New Albany, IN.
1. Prepared lab materials for department laboratory classes.
2. Assisted during laboratory sessions by fielding student queries.

1996 – 2006 **Boatswains Mate First Class (Enlisted Warfare Specialist)**. United States Navy, San Diego, CA and United States Naval Reserves, Louisville, KY.
1. Work center supervisor as a senior non-commissioned officer.
2. Performed safety officer duties during various high-risk operations.
3. Left occupation to pursue college degree full-time.

Honors, Awards, and Fellowships

2014 Received an American Society of Gene & Cell Therapy 17th annual meeting Meritorious Abstract Travel Award.
2014 Received Indiana University School of Medicine Spring Travel Award.
2012 Inducted into the SALUTE National Veterans Honor Society
2009 Received the Indiana University Southeast, School of Natural Sciences William B. Hebard Memorial Scholarship.
2008 Named Indiana University Southeast, School of Natural Sciences "2008 Outstanding Student of the Year in Biology".
2008 Awarded an Indiana University Southeast, School of Natural Sciences, Undergraduate Research Fellowship.
2006 Received the Summer Indiana University Southeast Misc. Scholarship

Presentations

2014 Oral presentation of "Optimizing Expression and Minimizing Integration in Non-Integrating Lentiviral Vectors" at the American Society of Gene and Cell Therapy 17th annual meeting in Washington, DC
2013 Departmental seminar series oral presentation "Safety Considerations in Lentiviral Gene Therapy" at Indiana University School of Medicine, Department of Medical and Molecular Genetics
2012 Poster presentation of graduate research "A Novel U3 LTR Integrase DNA Attachment Site Deleted Non-integrating Lentiviral Vector" at the American Society of Gene and Cell Therapy 15th annual meeting in Philadelphia, Pennsylvania
2011 Poster presentation of graduate research "Using Pulmozyme DNase Treatment as an Alternative to Benzonase Treatment for Lentiviral Supernatant Production" at the American Society of Gene and Cell Therapy 14th annual meeting in Seattle, Washington
2009 Oral presentation of research fellowship "Selenium Bioremediation with Macrofiber Forming *Bacillus Subtilis*" at the Indiana University Southeast Student Research Conference

Publications

Shaw, A. and Cornetta, K. Design and potential of non-integrating lentiviral vectors. Biomedicines 2: 14-35, 2014

Shaw, A., Bischof, D., Jasti, A., Ernstberger, A., Hawkins, T. and Cornetta, K. Using Pulmozyme DNase treatment in lentiviral vector production. Human Gene Therapy Methods: Part B 23: 65–71, 2012.

Profession Organizations

American Society of Gene & Cell Therapy (ASGCT), Associate Member 2011 – 2012, 2014